

OPTIMAL ROBUST MEAN-VARIANCE HEDGING IN INCOMPLETE FINANCIAL MARKETS

N. LAZRIEVA, T. TORONJADZE

Georgian–American University, Business School, 3 Alleyway II,
Chavchavadze Ave. 17 a, Tbilisi, Georgia, E-mail: toronj333@yahoo.com

A. Razmadze Mathematical Institute, 1, M. Aleksidze St., Tbilisi, Georgia

ABSTRACT. Optimal B -robust estimate is constructed for multidimensional parameter in drift coefficient of diffusion type process with small noise. Optimal mean-variance robust (optimal V -robust) trading strategy is found to hedge in mean-variance sense the contingent claim in incomplete financial market with arbitrary information structure and misspecified volatility of asset price, which is modelled by multidimensional continuous semimartingale. Obtained results are applied to stochastic volatility model, where the model of latent volatility process contains unknown multidimensional parameter in drift coefficient and small parameter in diffusion term.

1. INTRODUCTION, MOTIVATION AND RESULTS

The hedging and pricing of contingent claims in incomplete financial markets, and dynamic portfolio selection problems are important issues in modern theory of finance. These problems are associated due to the so-called mean-variance approach.

For determining a “good” hedging strategy in incomplete market with arbitrary information structure $F = (\mathcal{F})_{0 \leq t \leq T}$, one riskless asset and d , $d \geq 1$, risky assets, whose price process is a semimartingale X , the mean-variance approach suggests to use the quadratic criterion to measure the hedging error, i.e. to solve the mean-variance hedging problem introduced by Föllmer and Sondermann [10]:

$$\text{minimize } E \left(H - x - \int_0^T \theta_t dX_t \right)^2 \text{ over all } \theta \in \Theta, \quad (1.1)$$

where contingent claim H is a \mathcal{F}_T -measurable square-integrable random variable (r.v.), x is an initial investment, Θ is a class of admissible trading strategies, T is an investment horizon.

The mean-variance formulation by Markowitz [26], provides a foundation for a single period portfolio selection (see, also Merton [27]). In recent paper of Li

1991 *Mathematics Subject Classification.* 60G22, 62F35, 91B28, 62F35, 62M05, 62M09.

Key words and phrases. Stochastic volatility, small diffusion, robust parameter estimate, optimal mean-variance robust hedging.

and Ng [22] the concept of Markowitz's mean-variance formulation for finding the optimal portfolio policy and determining the efficient frontier in analytical form has been extended to multiperiod portfolio selection.

As it pointed out in Li and Ng [22] the results on multiperiod mean-variance formulation with one riskless asset can be derived using the results of the mean-variance hedging formulation.

Therefore, the mean-variance hedging is a powerful approach for both above mentioned major problems.

The problem (1.1) was intensively investigated in last decade (see, e.g., Duffie and Richardson [9], Schweizer [36], [37], [38], Delbaen et al. [8], Monat and Striker [28], Rheinländer and Schweizer [33], (RSch hereafter), Pham et al. [31], Gourieroux et al. [11] (GLP hereafter), Laurent and Pham [18]).

A stochastic volatility model, proposed by Hull and White [13] and Scott [39], where the stock price volatility is an random process, is a popular model of incomplete market, where the mean-variance hedging approach can be used (see, e.g., Laurent and Pham [18], Biagini et al. [13], Mania and Tevzadze [24], Pham et al. [31]).

Consider the stochastic volatility model described by the following system of SDE

$$\begin{aligned} dX_t &= X_t dR_t, \quad X_0 > 0, \\ dR_t &= \mu_t(R_t, Y_t) dt + \sigma_t dw_t^R, \quad R_0 = 0, \\ \sigma_t^2 &= f(Y_t), \\ dY_t &= a(t, Y_t; \alpha) dt + \varepsilon dw_t^\sigma, \quad Y_0 = 0, \end{aligned} \tag{1.2}$$

where $w = (w^R, w^\sigma)$ is a standard two-dimensional Wiener process, defined on complete probability space (Ω, \mathcal{F}, P) , $F^w = (\mathcal{F}_t^w)_{0 \leq t \leq T}$ is the P -augmentation of the natural filtration $\mathcal{F}_t^w = \sigma(w_s, 0 \leq s \leq t)$, $0 \leq t \leq T$, generated by w , $f(\cdot)$ is a continuous one-to-one positive locally bounded function (e.g., $f(x) = e^x$), $\alpha = (\alpha_1, \dots, \alpha_m)$, $m \geq 1$, is a vector of unknown parameters, and ε , $0 < \varepsilon \ll 1$, is a small number. Assume that the system (1.2) has an unique strong solution.

This model is analogous to the model proposed by Renault and Touzi [32] (RT hereafter). The principal difference is the presence of small parameter ε in our model, which due to the assumption that the volatility of randomly fluctuated volatility process is small (see, also Sircar and Papanicolau [40]). Thus assumption enables us to use the prices of trading options with short, nearest to the current time value maturities for volatility process filtration and parameter estimation purposes (see below). In contrast, RT [32] needs to assume that there exist trading derivatives with any (up to the infinity) maturities.

Important feature of the stochastic volatility models is that volatility process Y is unobservable (latent) process. To obtain explicit form of optimal trading strategy full knowledge of the model of the process Y is necessary and hence one needs to estimate the unknown parameter $\alpha = (\alpha_1, \dots, \alpha_m)$, $m \geq 1$.

A variety of estimation procedures are used, which involve either direct statistical analysis of the historical data or the use of implied volatilities extracted from prices of existing traded derivatives.

For example, one can use the following method based on historical data.

Fix the time variable t . From observations $X_{t_0^{(n)}}, \dots, X_{t_n^{(n)}}$, $0 = t_0^{(n)} < \dots < t_n^{(n)} = t$, $\max_j [t_{j+1}^{(n)} - t_j^{(n)}] \rightarrow 0$, as $n \rightarrow \infty$, calculate the realization of yield process $R_t = \int_0^t \frac{dX_s}{X_s}$, and then calculate the sum

$$S_n(t) = \sum_{j=0}^{n-1} |R_{t_{j+1}^{(n)}} - R_{t_j^{(n)}}|^2.$$

It is well-known (see, e.g., Lipster and Shiryaev [23]) that

$$S_n(t) \xrightarrow{P} \int_0^t \sigma_s^2 ds \quad \text{as } n \rightarrow \infty.$$

Since $\sigma_t^2(\omega) = f(Y_t)$ is a continuous process we get

$$\sigma_t^2(\omega) = \lim_{\Delta \downarrow 0} \frac{F(t + \Delta, \omega) - F(t, \omega)}{\Delta},$$

where $F(t, \omega) = \int_0^t \sigma_s^2(\omega) ds$.

Hence, the realization $(y_t)_{0 \leq t \leq T}$ of the process Y can be found by the formula $y_t = f^{-1}(\sigma_t^2)$, $0 \leq t \leq T$.

More sophisticated methods using the same idea can be found, e.g., in Chesney et al. [5], Pastorello [30].

We can use the reconstructed sample path (y_t) , $0 \leq t \leq T$, to estimate the unknown parameter α in the drift coefficient of diffusion process Y .

The second, market price adjusted procedure of reconstruction the sample path of volatility process Y and parameter estimate was suggested by RT [32], where they used implied volatility data.

We present a quick review of this method, adapted to our model (1.2).

Suppose that the volatility risk premium $\lambda^\sigma \equiv 0$, meaning that the risk from the volatility process is non-compensated (or can be diversified away). Then the price $C_t(\sigma)$ of European call option can be calculated by the Hull and White formula (see, e.g., RT [32]), and Black-Scholes (BS) implied volatility $\sigma^i(\sigma)$ can be found as a unique solution of the equation

$$C_t(\sigma) = C_t^{BS}(\sigma^i(\sigma)),$$

where $C^{BS}(\sigma)$ denotes the standard BS formula written as a function of the volatility parameter σ .

Here (for further estimational purposes) only at-the-money options are used.

Under some technical assumptions (see Proposition 5.1 of RT [32], and Buejeux and Rochet [23] for general diffusion of volatility process)

$$\frac{\partial \sigma_t^i(\sigma, \alpha)}{\partial \sigma_t} > 0 \quad (1.3)$$

(remember that the drift coefficient of process Y depends on unknown parameter α).

Fix current value of time parameter t , $0 \leq t \leq T$, and let $0 < T_1 < T_2 < \dots < T_{k-1} < t < T_k$ be the maturity times of some traded at-the-money options.

Let $\sigma_{t_j^\varepsilon}^{i*}$ be the observations of an implied volatility at the time moments $0 = t_0^\varepsilon < t_1^\varepsilon < \dots < t_{[\frac{t}{\varepsilon}]}^\varepsilon = t$, $\max_j [t_{j+1}^\varepsilon - t_j^\varepsilon] \rightarrow 0$, as $\varepsilon \rightarrow 0$.

Then, using (1.3), and solving the equation

$$\sigma_{t_j^\varepsilon}^i(\sigma_{t_j^\varepsilon}, \alpha) = \sigma_{t_j^\varepsilon}^{i*},$$

one can obtain the realization $\{\tilde{\sigma}_{t_j^\varepsilon}\}$ of the volatility (σ_t) , and thus, using the formula $y_{t_j^\varepsilon} = f^{-1}(\tilde{\sigma}_{t_j^\varepsilon}^2)$, the realization $\{y_{t_j^\varepsilon}\}$ of volatility process (Y_t) , which can be viewed as the realization of nonlinear AR(1) process:

$$Y_{t_{j+1}^\varepsilon} - Y_{t_j^\varepsilon} = a(t_j^\varepsilon, Y_{t_j^\varepsilon}; \alpha)(t_{j+1}^\varepsilon - t_j^\varepsilon) + \varepsilon(w_{t_{j+1}^\varepsilon}^\sigma - w_{t_j^\varepsilon}^\sigma).$$

Using the data $\{y_{t_j^\varepsilon}\}$ one can construct the MLE $\hat{\alpha}_t^\varepsilon$ of parameter α , see, e.g., Chitashvili et al. [25], [26], Lazrieva and Toronjadze [19].

Remember the scheme of construction of MLE. Rewrite the previous AR(1) process, using obvious simple notation, in form

$$Y_{j+1} - Y_j = a(t_j, Y_j; \alpha)\Delta + \varepsilon\Delta w_j^\sigma.$$

Then

$$\begin{aligned} \frac{\partial}{\partial y} P\{Y_{j+1} \leq y \mid Y_j\} &= \frac{1}{\sqrt{2\pi\Delta\varepsilon}} \exp\left(-\frac{(y - Y_j - a(t_j, Y_j; \alpha)\Delta)^2}{2\varepsilon^2\Delta}\right) \\ &=: \varphi_{j+1}(y, Y_j; \alpha), \end{aligned}$$

and the log-derivative of the likelihood process $\ell_t = (\ell_t^{(1)}, \dots, \ell_t^{(m)})$ is given by the relation

$$\ell_t^{(i)} = \sum_j \ell_{j+1}^{(i)}, \quad i = \overline{1, m},$$

where

$$\begin{aligned} \ell_{j+1}^{(i)}(y; \alpha) &= \frac{\partial}{\partial \alpha_i} \ln \varphi_{j+1}(y, Y_j; \alpha) \\ &= \frac{1}{\varepsilon^2\Delta} (y - Y_j - a(t_j, Y_j; \alpha)\Delta) \dot{a}^{(i)}(t_j, Y_j; \alpha)\Delta. \end{aligned}$$

Hence MLE is a solution (under some conditions) of the system of equations

$$\frac{1}{\varepsilon^2\Delta} \sum_j (y_{j+1} - y_j - a(t_j, y_j; \alpha)\Delta) \dot{a}^{(i)}(t_j, y_j; \alpha)\Delta = 0, \quad i = \overline{1, m},$$

where the reconstructed data $\{y_j\} = \{y_{t_j^\varepsilon}\}$ are substituted).

Following RT [32] let us introduce the functionals

$$\begin{aligned} HW_\varepsilon^{-1} : \hat{\alpha}_t^\varepsilon(p) &\rightarrow \left(y_{t_j^\varepsilon}^{(p+1)}, \quad 0 \leq j \leq \left\lfloor \frac{t}{\varepsilon} \right\rfloor \right), \\ MLE_\varepsilon : \left(y_{t_j^\varepsilon}^{(p+1)}, \quad 0 \leq j \leq \left\lfloor \frac{t}{\varepsilon} \right\rfloor \right) &\rightarrow \hat{\alpha}_t^\varepsilon(p+1) \end{aligned}$$

and

$$\phi_\varepsilon = MLE_\varepsilon \circ HW_\varepsilon^{-1}.$$

Starting with some constant initial value (or preliminary estimate obtained, e.g., from historical data) one can compute a sequence of estimates

$$\hat{\alpha}_t^\varepsilon(p+1) = \phi_\varepsilon(\hat{\alpha}_t^\varepsilon(p)), \quad p \geq 1.$$

If the operator ϕ_ε is a strong contraction in the neighborhood of the true value of the parameter α^0 , for a small enough ε , then one can define the estimate $\hat{\alpha}_t^\varepsilon$ as the limits of the sequence $\{\hat{\alpha}_t^\varepsilon(p)\}_{p \geq 1}$. It was proved in RT [32] that $\hat{\alpha}_t^\varepsilon$ is a strong consistent estimate of the parameter α .

Return to our consideration.

Interpolating on some way the corresponding (to the estimate $\hat{\alpha}_t^\varepsilon$) realization $\{y_{t_j^\varepsilon}\}$ we get the reconstructed continuous sample path $(y_s)_{0 \leq s \leq t}$ of the latent process Y , which can be used for further analysis.

Unfortunately, both described statistical procedures are highly sensitive w.r.t errors in all steps of parameter identification process.

Hence, this is a natural place for introducing the robust procedure of parameter estimates.

Suppose that the sample path $(y_s)_{0 \leq s \leq t}$ comes from the observation of process $(\tilde{Y}_s)_{0 \leq s \leq t}$ with distribution $\tilde{P}_\alpha^\varepsilon$ from the shrinking contamination neighborhood of the distribution P_α^ε of the basic process $Y = (Y_s)_{0 \leq s \leq t}$. That is

$$\frac{d\tilde{P}_\alpha^\varepsilon}{dP_\alpha^\varepsilon} \Big| \mathcal{F}_t^w = \mathcal{E}_t(\varepsilon N^\varepsilon), \quad (1.4)$$

where $N^\varepsilon = (N_s^\varepsilon)_{0 \leq s \leq t}$ is a P_α^ε -square integrable martingale, $\mathcal{E}_t(M)$ is the Dolean exponential of martingale M .

In the diffusion-type framework (1.4) represents the Huber gross error model (as it explain in Remark 2.2). The model of type (1.4) of contamination of measures for statistical models with filtration was suggested by Lazrieva and Toronjadze [20], [21].

In Section 2 we study the problem of construction of robust estimates for contamination model (1.4).

In subsection 2.1 we give a description of the basic model and definition of consistent uniformly linear asymptotically normal (CULAN) estimates, connected with the basic model (Definition 2.1).

In subsection 2.2 we introduce a notion of shrinking contamination neighborhood, described in terms of contamination of nominal distribution, which naturally leads to the class of alternative measures (see (2.18) and (2.19)).

In subsection 2.3 we study the asymptotic behaviour of CULAN estimates under alternative measures (Proposition 2.2), which is the basis for the formulation of the optimization problem.

In subsection 2.4 the optimization problem is solved which leads to construction of optimal B -robust estimate (Theorem 2.1).

Based on the limit theorem (subsection 2.1), one can construct the asymptotic confidence region of level γ for unknown parameter α

$$\lim_{\varepsilon \rightarrow \infty} P_{\alpha}^{\varepsilon}(\varepsilon^{-2}(\alpha - \alpha_t^{*,\varepsilon})'V^{-1}(\psi^*; \alpha_t^{*,\varepsilon})(\alpha - \alpha_t^{*,\varepsilon}) \leq \chi_{\gamma}^2) = 1 - \gamma,$$

where χ_{γ}^2 is a quantile of order $1 - \gamma$ of χ^2 -distribution with m degree of freedom, and $V(\psi^*; \alpha)$ is given by (2.17).

This region shrinks to the estimate $\alpha_t^{*,0}$, as $\varepsilon \rightarrow 0$.

Now if the coefficient $a(t, y; \alpha)$ in (1.2) is such that the solution $Y_t^{\varepsilon}(\alpha)$ of SDE (1.2) is continuous w.r.t parameter α (see, e.g., Krylov [16]), then the confidence region of parameter α is mapped to the confidence interval for $Y_t^{\varepsilon}(\alpha)$, which shrinks to $Y_t^* = Y_t^0(\alpha_t^{*,0})$. Further, by the function f , the latter interval is mapped to the confidence interval for σ_t , which shrinks to $\sigma_t^* = f^{1/2}(Y_t^0(\alpha_t^{*,0}))$. Denote σ_t^0 the center of this interval. Then the interval can be written in the form

$$\sigma_t = \sigma_t^0 + \delta(\varepsilon)h_t,$$

where $\delta(\varepsilon) \rightarrow 0$, as $\varepsilon \rightarrow 0$, and $h \in \mathcal{H}$ (see (3.18)).

Thus, we arrive at the asset price model (1.2) with misspecified volatility, and it is natural to consider the problem of construction of the robust trading strategy to hedge a contingent claim H .

We investigate this problem in the mean-variance setting in Section 3. We consider the general situation, when the asset price is modelled by d -dimensional continuous semimartingale and the information structure is given by some general filtration.

In subsection 3.1 we give a description of the financial market model.

In subsection 3.2 we collect the facts concerning the variance-optimal equivalent local martingale measure, which plays a key role in the mean-variance hedging approach.

In last subsection 3.3 we construct “optimal robust hedging strategy” (Theorem 3.1) by approximating the optimization problem (3.25) by the problem (3.27). As it is mentioned in Remark 3.2, such approach and term are common in robust statistic theory. In contact to optimal B -robustness (see Section 2), here we develop the approach, known in robust statistics as optimal V -robustness, see Hampel et al. [12].

Note that our approach allows incorporating current information on the underlying model, and hence is adaptive. Namely, passing from time value t to $t + \tau$, $\tau > 0$, when more information about market prices are available,

the asymptotic variance-covariance of the constructed estimate $\alpha_t^{*,\varepsilon}$ becomes smaller, and hence the estimation procedure becomes more precise.

In the paper of Runggaldier and Zaccaria [35] the adaptive approach to risk management under general uncertainty (restricted information) was developed. As it is mentioned in this paper there exist a series of investigations dealt with various type of adaptive approaches (see list of references in [35]). But in all these papers (except Runggaldier and Zaccaria [35]) the uncertainty is only in the stock appreciation rate in contrast to our consideration, where the model misspecification is due to the volatility parameter.

The consideration of misspecified asset price models was initiated by Avellaneda et al. [1], Avellaneda and Paras [2].

Various authors in different settings attacked the robustness problem. The method used in Section 3 was suggested by Toronjadze [41] for asset price process modelled by the one-dimensional process. As it will be shown in Remark 3.2 below, in simplest case when asset price process is a martingale w.r.t initial measure P , and it is possible to find the solution of “exact” optimization problem (3.25), this solution coincides with the solution of an approximating optimization problem (3.27). In more general situation (when asset price process is not more the P -martingale) investigation of the problem (3.25) by, e.g., control theory methods seems sufficiently difficult. Anyway, we do not know the solution of the problem (3.25).

Return to the stochastic volatility model (1.2) and describe successive steps of our approach:

- 1) For each current time value t , $0 < t < T$, reconstruct the sample path $(y_s)_{0 \leq s \leq t}$, using the historical data or the tradable derivatives prices;
- 2) Using the approach developed in Section 2, calculate the value $\alpha_t^{*,\varepsilon}$ of the robust estimate of parameter α (i.e. construct the deterministic function $t \rightarrow \alpha_t^{*,\varepsilon} \in R^m$) and then find the confidence region for parameter α ;
- 3) Based on the volatility process model find the confidence interval for $Y_t(\alpha)$;
- 4) Denoting $a^*(t, y) = a(t, y; \alpha_t^{*,\varepsilon})$, where $a(t, y; \alpha)$ is a drift coefficient of volatility process, consider the stochastic volatility model with misspecified asset price model and fully specified volatility process model

$$\begin{aligned} dX_t &= X_t dR_t, \quad X_0 > 0, \\ dR_t &= (\sigma_t^0 + \delta(\varepsilon)h_t)dM_t^0, \quad R_0 = 0, \\ dY_t &= a^*(t, Y_t)dt + \varepsilon dw_t^\sigma, \quad Y_0 = 0, \quad 0 \leq t \leq T, \end{aligned}$$

where

$$dM_t^0 = k_t dt + dw_t^R,$$

$h \in \mathcal{H}$ and σ_t^0 is the center of the confidence interval of volatility.

Using Theorem 3.1 construct the optimal robust hedging strategy by the formula (3.44),

$$\theta_t^* = \frac{1}{\sigma_t^0} \left[\psi_t^{1,H} + \zeta_t(V_t^* - (\psi_t^H)'U_t) \right],$$

where all objects are defined in Theorem 3.1. \square

It should be mentioned that if one constructs a hedging strategy $\tilde{\theta}_t^*$ by the above-given formula with $\sigma_t^{*,\varepsilon} = f^{1/2}(Y_t^\varepsilon(\alpha_t^{*,\varepsilon}))$ instead of σ_t^0 , then the strategies $\tilde{\theta}_t^*$ and θ_t^* would be different, since $\sigma_t^{*,\varepsilon} \neq \sigma_t^0$, in general. Hence the value $\Delta_t = |\sigma_t^{*,\varepsilon} - \sigma_t^0|$ defines the correction term between the robust, θ_t^* and non-robust, $\tilde{\theta}_t^*$ strategies.

In nontrivial case, when $k_t = k(Y_t)$ the variance-optimal martingale measure \tilde{P} is given by (3.17), $\zeta_t = -k_t \mathcal{E}_t(-k \cdot M^0)$ (see subsection 3.2), and the process $(X_t, Y_t)_{0 \leq t \leq T}$ is the Markov process. If $H = h(X_T, Y_T)$ ($h(x, y)$ is some function), then $\tilde{V}_t^H = E^{\tilde{P}}(H | \mathcal{F}_t^w) = E^{\tilde{P}}(h(X_T, Y_T) | \mathcal{F}_t^w) = v(t, X_t, Y_t)$ and if, e.g., $v(t, x, y) \in C^{1,2,2}$, then v is a unique solution of the following partial differential equation

$$\frac{\partial v}{\partial t} + a^* \frac{\partial v}{\partial y} + \frac{1}{2} \left(\varepsilon^2 \frac{\partial^2 v}{\partial y^2} + x^2 v^2 \frac{\partial^2 v}{\partial x^2} \right) = 0,$$

with the boundary condition $v(T, x, t) = h(x, y)$. More general situation with nonsmooth v is considered in Laurent and Pham [18], Mania and Tevzadze [24].

Further, one can find the Galtchouk–Kunita–Watanabe decomposition of r.v. H (see, e.g., Pham et al. [31]) putting

$$\xi_t^H = \frac{\partial v(t, X_t, Y_t)}{\partial x}, \quad L_T^H = \varepsilon \int_0^T \frac{\partial v}{\partial y}(t, X_t, Y_t) dw_t^\sigma,$$

and calculate ψ_t^H , L_T and V_t^* using (4.13) and (4.14) of RSch [33].

Thus one gets the explicit solution of the mean-variance hedging problem.

Finally, here is the short summary of approach:

a) Incorporate the robust procedure in statistical analysis of volatility process. That is construct and use in the model optimal B -robust estimate of unknown parameter in drift coefficient of volatility process.

Parameter estimation naturally leads to the asset price model misspecification.

b) Incorporate the second robust procedure in financial analysis of contingent claim hedging. That is construct and use for hedging purposes optimal V -robust trading strategy.

In our opinion this “double robust” strategy should be more attractive to protect the hedger against the possible errors.

The general asymptotic theory of estimation can be found in Ibragimov and Khas'miskii [14]; the theory of robust statistics is developed in Hampel et al. [12] and in Rieder [34]; the theory of the trend parameter estimates for diffusion process with small noise is developed in Kutoyants [17]; the book of Musiela and Rutkowski [29] is devoted to the mathematical theory of finance and finally, the general theory of martingales can be found in Jacod and Shiryaev [15].

2. OPTIMAL B -ROBUST ESTIMATES

2.1. Basic model. CULAN estimates. The basic model of observations is described by the SDE

$$dY_s = a(s, Y; \alpha) ds + \varepsilon dw_s, \quad Y_0 = 0, \quad 0 \leq s \leq t, \quad (2.1)$$

where t is a fixed number, $w = (w_s)_{0 \leq s \leq t}$ is a standard Wiener process defined on the filtered probability space $(\Omega, \mathcal{F}, F = (\mathcal{F}_s)_{0 \leq s \leq t}, P)$ satisfying the usual conditions, $\alpha = (\alpha_1, \dots, \alpha_m)$, $m \geq 1$, is an unknown parameter to be estimated, $\alpha \in \mathcal{A} \subset R^m$, \mathcal{A} is an open subset of R^m , ε , $0 < \varepsilon \ll 1$, is a small parameter (index of series). In our further considerations all limits correspond to $\varepsilon \rightarrow 0$.

Denote (C_t, \mathcal{B}_t) a measurable space of continuous on $[0, t]$ functions $x = (x_s)_{0 \leq s \leq t}$ with σ -algebra $\mathcal{B}_t = \sigma(x : x_s, s \leq t)$. Put $\mathcal{B}_s = \sigma(x : x_u, u \leq s)$.

Assume that for each $\alpha \in \mathcal{A}$ the drift coefficients $a(s, x; \alpha)$, $0 \leq s \leq t$, $x \in C_t$ is a known nonanticipative (i.e. \mathcal{B}_s -measurable for each s , $0 \leq s \leq t$) functional satisfying the functional Lipschitz and linear growth conditions **L**:

$$|a(s, x^1; \alpha) - a(s, x^2; \alpha)| \leq L_1 \int_0^s |x_u^1 - x_u^2| dk_u + L_2 |x_s^1 - x_s^2|,$$

$$|a(s, x; \alpha)| \leq L_1 \int_0^s (1 + |x_u|) dk_u + L_2 (1 + |x_s|),$$

where L_1 and L_2 are constants, which do not depend on α , $k = (k(s))_{0 \leq s \leq t}$ is a non-decreasing right-continuous function, $0 \leq k(s) \leq k_0$, $0 : k_0 < \infty$, $x^1, x^2 \in C_t$.

Then, as it is well-known (see, e.g., Liptser and Shiryaev [23]), for each $\alpha \in \mathcal{A}$ the equation (2.1) has an unique strong solution $Y^\varepsilon(\alpha) = (Y_s^\varepsilon(\alpha))_{0 \leq s \leq t}$, and in addition (see Kutoyants [17])

$$\sup_{0 \leq s \leq t} |Y_s^\varepsilon(\alpha) - Y_s^0(\alpha)| \leq C\varepsilon \sup_{0 \leq s \leq t} |w_s| \quad P\text{-a.s.},$$

with some constant $C = C(L_1, L_2, k_0, t)$, where $Y^0(\alpha) = (Y_s^0(\alpha))_{0 \leq s \leq t}$ is the solution of the following nonperturbed differential equation

$$dY_s = a(s, Y; \alpha) ds, \quad Y_0 = 0. \quad (2.2)$$

Change the initial problem of estimation of parameter α by the equivalent one, when the observations are modelled according to the following SDE

$$dX_s = a_\varepsilon(s, X; \alpha) ds + dw_s, \quad X_0 = 0, \quad (2.3)$$

where $a_\varepsilon(s, x; \alpha) = \frac{1}{\varepsilon} a(s, \varepsilon x; \alpha)$, $0 \leq s \leq t$, $x \in C_t$, $\alpha \in \mathcal{A}$.

It is clear that if $X^\varepsilon(\alpha) = (X_s^\varepsilon(\alpha))_{0 \leq s \leq t}$ is the solution of SDE (2.3), then for each $s \in [0, t]$ $\varepsilon X_s^\varepsilon(\alpha) = Y_s^\varepsilon(\alpha)$.

Denote by P_α^ε the distribution of process $X^\varepsilon(\alpha)$ on the space (C_t, \mathcal{B}_t) , i.e. P_α^ε is the probability measure on (C_t, \mathcal{B}_t) induced by the process $X^\varepsilon(\alpha)$. Let P^w

be a Wiener measure on (C_t, \mathcal{B}_t) . Denote $X = (X_s)_{0 \leq s \leq t}$ a coordinate process on (C_t, \mathcal{B}_t) , that is $X_s(x) = x_s$, $x \in C_t$.

The conditions **L** guarantee that for each $\alpha \in \mathcal{A}$ the measures P_α^ε and P^w are equivalent ($P_\alpha^\varepsilon \sim P^w$), and if we denote $z_s^{\alpha, \varepsilon} = \frac{dP_\alpha^\varepsilon}{dP^w} | \mathcal{B}_s$ the density process (likelihood ratio process), then

$$z_s^{\alpha, \varepsilon}(X) = \mathcal{E}_s(a_\varepsilon(\alpha) \cdot X) := \exp \left\{ \int_0^s a_\varepsilon(u, X; \alpha) dX_u - \frac{1}{2} \int_0^s a_\varepsilon^2(u, X; \alpha) du \right\}.$$

Introduce class Ψ of R^m -valued nonanticipative functionals ψ , $\psi : [0, t] \times C_t \times \mathcal{A} \rightarrow R^m$ such that for each $\alpha \in \mathcal{A}$ and $\varepsilon > 0$

$$1) \quad E_\alpha^\varepsilon \int_0^t |\psi(s, X; \alpha)|^2 ds < \infty, \quad (2.4)$$

$$2) \quad \int_0^t |\psi(s, Y^0(\alpha); \alpha)|^2 ds < \infty, \quad (2.5)$$

3) uniformly in α on each compact $K \subset \mathcal{A}$

$$P_\alpha^\varepsilon - \lim_{\varepsilon \rightarrow 0} \int_0^t |\psi(s, \varepsilon X; \alpha) - \psi(s, Y^0(\alpha); \alpha)|^2 ds = 0, \quad (2.6)$$

where $|\cdot|$ is an Euclidean norm in R^m , $P_\alpha^\varepsilon - \lim_{\varepsilon \rightarrow 0} \zeta_\varepsilon = \zeta$ denotes the convergence $P_\alpha^\varepsilon \{|\zeta_\varepsilon - \zeta| > \rho\} \rightarrow 0$, as $\varepsilon \rightarrow 0$, for all ρ , $\rho > 0$.

Assume that for each $s \in [0, t]$ and $x \in C_t$ the functional $a(s, x; \alpha)$ is differentiable in α and gradient $\dot{a} = \left(\frac{\partial}{\partial \alpha_1} a, \dots, \frac{\partial}{\partial \alpha_m} a \right)'$ belongs to Ψ ($\dot{a} \in \Psi$), where the sign “'” denoted a transposition.

Then the Fisher information process

$$I_s^\varepsilon(X; \alpha) := \int_0^s \dot{a}_\varepsilon(u, X; \alpha) [\dot{a}_\varepsilon(u, X; \alpha)]' du, \quad 0 \leq s \leq t,$$

is well-defined and, moreover, uniformly in α on each compact

$$P_\alpha^\varepsilon - \lim_{\varepsilon \rightarrow 0} \varepsilon^2 I_t^\varepsilon(\alpha) = I^0(\alpha), \quad (2.7)$$

where

$$I^0(\alpha) := \int_0^t \dot{a}(s, Y^0(\alpha); \alpha) [\dot{a}(s, Y^0(\alpha); \alpha)]' ds.$$

For each $\psi \in \Psi$, introduce the functional $\psi_\varepsilon(s, x; \alpha) := \frac{1}{\varepsilon} \psi(s, \varepsilon x; \alpha)$ and matrices $\Gamma_\varepsilon^\psi(\alpha)$ and $\gamma_\varepsilon^\psi \alpha$:

$$\Gamma_\varepsilon^\psi(X; \alpha) := \int_0^t \psi_\varepsilon(s, X; \alpha) [\psi_\varepsilon(s, X; \alpha)]' ds, \quad (2.8)$$

$$\gamma_\varepsilon^\psi(X; \alpha) := \int_0^t \psi_\varepsilon(s, X; \alpha) [\dot{a}_\varepsilon(s, X; \alpha)]' ds. \quad (2.9)$$

Then from (2.6) it follows that uniformly in α on each compact

$$P_\alpha^\varepsilon - \lim_{\varepsilon \rightarrow 0} \varepsilon^2 \Gamma_\varepsilon^\psi(\alpha) = \Gamma_0^\psi(\alpha), \quad (2.10)$$

$$P_\alpha^\varepsilon - \lim_{\varepsilon \rightarrow 0} \varepsilon^2 \gamma_\varepsilon^\psi(\alpha) = \gamma_0^\psi(\alpha), \quad (2.11)$$

where the matrices $\Gamma_0^\psi(\alpha)$ and $\gamma_0^\psi(\alpha)$ are defined as follows

$$\Gamma_0^\psi(\alpha) := \int_0^t \psi(s, Y^0(\alpha); \alpha) [\psi(s, Y^0(\alpha); \alpha)]' ds, \quad (2.12)$$

$$\gamma_0^\psi(\alpha) := \int_0^t \psi(s, Y^0(\alpha); \alpha) [\dot{a}(s, Y^0(\alpha); \alpha)]' ds. \quad (2.13)$$

Note that, by virtue of (2.4), (2.5) and $\dot{a} \in \Psi$, matrices given by (2.8), (2.9), (2.12) and (2.13) are well-defined.

Denote Ψ_0 the subset of Ψ such that for each $\psi \in \Psi_0$ and $\alpha \in \mathcal{A}$, $\text{rank } \Gamma_0^\psi(\alpha) = m$ and $\text{rank } \gamma_0^\psi(\alpha) = m$.

Assume that $\dot{a} \in \Psi_0$.

For each $\psi \in \Psi_0$, define a P_α^ε -square integrable martingale $L^{\psi, \varepsilon}(\alpha) = (L_s^{\psi, \varepsilon}(\alpha))_{0 \leq s \leq t}$ as follows

$$L_s^{\psi, \varepsilon}(X; \alpha) = \int_0^s \psi_\varepsilon(u, X; \alpha) (dX_u - \alpha_\varepsilon(u, X; \alpha) du). \quad (2.14)$$

Now we give a definition of CULAN M -estimates.

Definition 2.1. An estimate $(\alpha_t^{\psi, \varepsilon})_{\varepsilon > 0} = (\alpha_{1,t}^{\psi, \varepsilon}, \dots, \alpha_{m,t}^{\psi, \varepsilon})'_{\varepsilon > 0}$, $\psi \in \Psi_0$, is called consistent uniformly lineal asymptotically normal (CULAN) if it admits the following expansion

$$\alpha_t^{\psi, \varepsilon} = \alpha + [\gamma_0^\psi(\alpha)]^{-1} \varepsilon^2 L_t^{\psi, \varepsilon}(\alpha) + r_{\psi, \varepsilon}(\alpha), \quad (2.15)$$

where uniformly in α on each compact

$$P_\alpha^\varepsilon - \lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} r_{\psi, \varepsilon}(\alpha) = 0. \quad (2.16)$$

It is well-known (see Lazrieva, Toronjadze [19]) that under the above conditions uniformly in α on each compact

$$\mathcal{L}\{\varepsilon^{-1}(\alpha_t^{\psi,\varepsilon} - \alpha) \mid P_\alpha^\varepsilon\} \xrightarrow{w} N(0, V(\psi; \alpha)),$$

with

$$V(\psi; \alpha) := [\gamma_0^\psi(\alpha)]^{-1} \Gamma_0^\psi(\alpha) ([\gamma_0^\psi(\alpha)]^{-1})', \quad (2.17)$$

where $\mathcal{L}(\zeta|P)$ denotes the distribution of random vector ζ calculated under measure P , symbol “ \xrightarrow{w} ” denotes the weak convergence of measures, $N(0, V(\psi; \alpha))$ is a distribution of Gaussian vector with zero mean and covariance matrix $V(\psi; \alpha)$.

Remark 2.1. In context of diffusion type processes the M -estimate $(\alpha_t^{\psi,\varepsilon})_{\varepsilon>0}$ is defined as a solution of the following stochastic equation

$$L_t^{\psi,\varepsilon}(X; \alpha) = 0,$$

where $L_t^{\psi,\varepsilon}(X; \alpha)$ is defined by (2.14), $\psi \in \Psi_0$.

The asymptotic theory of M -estimates for general statistical models with filtration is developed in Chitashvili et al. [7]. Namely, the problem of existence and global behaviour of solutions is studied. In particular, the conditions of regularity and ergodicity type are established, under which M -estimates have a CULAN property.

For our model, in case when $\mathcal{A} = R^m$, the sufficient conditions for CULAN property take the form:

- (1) for all s , $0 \leq s \leq t$, and $x \in C_t$ the functionals $\psi(s, x; \alpha)$ and $\dot{a}(s, x; \alpha)$ are twice continuously differentiable in α with bounded derivatives satisfying the functional Lipschitz conditions with constants, which do not depend on α .
- (2) the equation (w.r.t y)

$$\Delta(\alpha, y) := \int_0^t \psi(s, Y^0(\alpha); y) (a(s, Y^0(\alpha); \alpha) - a(s, Y^0(\alpha); y)) ds = 0,$$

has an unique solution $y = \alpha$.

The MLE is a special case of M -estimates when $\psi = \dot{a}$.

Remark 2.2. According to (2.7) the asymptotic covariance matrix of MLE $(\hat{\alpha}_t^\varepsilon)_{\varepsilon>0}$ is $[I_0(\alpha)]^{-1}$. By the usual technique one can show that for each $\alpha \in \mathcal{A}$ and $\psi \in \Psi_0$, $I_0^{-1}(\alpha) \leq V(\psi, \alpha)$ (see (2.17)), where for two symmetric matrices B and C the relation $B \leq C$ means that the matrix $C - B$ is nonnegative definite.

Thus, the MLE has a minimal covariance matrix among all M -estimates.

2.2. Shrinking contamination neighborhoods. In this subsection we give a notion of a contamination of the basic model (2.3), described in terms of shrinking neighborhoods of basic measures $\{P_\alpha^\varepsilon, \alpha \in \mathcal{A}, \varepsilon > 0\}$, which is an analog of the Huber gross error model (see, e.g., Hampel et.al. [12] and also, Remark 2.3 below).

Let \mathcal{H} be a family of bounded nonanticipative functionals $h : [0, t] \times C_t \times \mathcal{A} \rightarrow R^1$ such that for all $s \in [0, t]$ and $\alpha \in \mathcal{A}$ the functional $h(s, x; \alpha)$ is continuous at the point $x_0 = Y^0(\alpha)$.

Let for each $h \in \mathcal{H}$, $\alpha \in \mathcal{A}$ and $\varepsilon > 0$, $P_\alpha^{\varepsilon, h}$ be a measure on (C_t, \mathcal{B}_t) such that

$$\begin{aligned} 1) \quad & P_\alpha^{\varepsilon, h} \sim P_\alpha^\varepsilon, \\ 2) \quad & \frac{dP_\alpha^{\varepsilon, h}}{dP_\alpha^\varepsilon} = \mathcal{E}_t(\varepsilon N_\alpha^{\varepsilon, h}), \end{aligned} \quad (2.18)$$

where

$$3) \quad N_{\alpha, s}^{\varepsilon, h} := \int_0^s h_s(u, X; \alpha)(dX_u - a_\varepsilon(u, X; \alpha) du), \quad (2.19)$$

with $h_\varepsilon(s, x; \alpha) := \frac{1}{\varepsilon} h(s, \varepsilon x; \alpha)$, $0 \leq s \leq t$, $x \in C_t$.

Denote $\mathbf{P}_\alpha^{\varepsilon, \mathcal{H}}$ a class of measures $P_\alpha^{\varepsilon, h}$, $h \in \mathcal{H}$, that is

$$\mathbf{P}_\alpha^{\varepsilon, \mathcal{H}} = \{P_\alpha^{\varepsilon, h}; \quad h \in \mathcal{H}\}.$$

We call $(\mathbf{P}_\alpha^{\varepsilon, \mathcal{H}})_{\varepsilon > 0}$ a shrinking contamination neighborhoods of the basic measures $(P_\alpha^\varepsilon)_{\varepsilon > 0}$, and the element $(P_\alpha^{\varepsilon, h})_{\varepsilon > 0}$ of these neighborhoods is called alternative measure (or simply alternative).

Obviously for each $h \in \mathcal{H}$ and $\alpha \in \mathcal{A}$, the process $N_\alpha^{\varepsilon, h} = (N_{\alpha, s}^{\varepsilon, h})_{0 \leq s \leq t}$ defined by (2.19) is a P_α^ε -square integrable martingale. Since under measure P_α^ε the process $\bar{w} = (\bar{w}_s)_{0 \leq s \leq t}$ defined as

$$\bar{w}_s := X_s - \int_0^s a_\varepsilon(u, X; \alpha) du, \quad 0 \leq s \leq t,$$

is a Wiener process, then by virtue of the Girsanov Theorem the process $\tilde{w} := \bar{w} + \langle \bar{w}, \varepsilon N_\alpha^{\varepsilon, h} \rangle$ is a Wiener process under changed measure $P_\alpha^{\varepsilon, h}$. But by the definition

$$\tilde{w}_s = X_s - \int_0^s (a_\varepsilon(u, X; \alpha) + \varepsilon h_\varepsilon(u, X; \alpha)) du,$$

and hence, one can conclude that $P_\alpha^{\varepsilon, h}$ is a weak solution of SDE

$$dX_s = (a_\varepsilon(s, X; \alpha) + \varepsilon h_\varepsilon(s, X; \alpha)) ds + dw_s, \quad X_0 = 0.$$

This SDE can be viewed as a “small” perturbation of the basic model (2.3).

Remark 2.3. 1) In the case of i.i.d. observations X_1, X_2, \dots, X_n , $n \geq 1$, the Huber gross error model in shrinking setting is defined as follows

$$f^{n, h}(x; \alpha) := (1 - \varepsilon_n) f(x; \alpha) + \varepsilon_n h(x; \alpha),$$

where $f(x; \alpha)$ is a basic (core) density of distribution of r.v. X_i (w.r.t some dominating measure μ), $h(x; \alpha)$ is a contaminating density, $f^{n, h}(x; \alpha)$ is a contaminated density, $\varepsilon_n = O(n^{-1/2})$. If we denote by P_α^n and $P_\alpha^{n, h}$ the measures

on $(R^n, \mathcal{B}(R^n))$, generated by $f(x; \alpha)$ and $f^{n,h}(x; \alpha)$, respectively, then

$$\frac{dP_\alpha^{n,h}}{dP_\alpha^n} = \prod_{i=1}^n \frac{f^{n,h}(X_i; \alpha)}{f(X_i; \alpha)} = \prod_{i=1}^n (1 + \varepsilon_n H(X_i; \alpha)) = \mathcal{E}_n(\varepsilon_n \cdot N_\alpha^{n,h}),$$

where $H = \frac{h-f}{f}$, $N_\alpha^{n,h} = (N_{\alpha,m}^{n,h})_{1 \leq m \leq n}$, $N_{\alpha,m}^{n,h} = \sum_{i=1}^m H(X_i; \alpha)$, $N_\alpha^{n,h}$ is a P_α^n -martingale, $\mathcal{E}_n(\varepsilon_n N_\alpha^{n,h}) = \prod_{i=1}^n (1 + \varepsilon_n \Delta N_{\alpha,i}^{n,h})$ is the Dolean exponential in discrete time case.

Thus

$$\frac{dP_\alpha^{n,h}}{dP_\alpha^n} = \mathcal{E}(\varepsilon_n \cdot N_\alpha^{n,h}), \quad (2.20)$$

and the relation (2.18) is a direct analog of (2.20).

2) The concept of shrinking contamination neighborhoods, expressed in the form of (2.18) was proposed in Lazrieva and Toronjadze [20] for more general situation, concerning with the contamination areas for semimartingale statistical models with filtration. \square

Note here that the power of the small parameter ε is crucial. One cannot consider the perturbation of measure with different power of ε if he/she wish to get nontrivial result.

In the remainder of this subsection we study the asymptotic properties of CULAN estimates under alternatives.

For this aim we first consider the problem of contiguity of measures $(P_\alpha^{\varepsilon,h})_{\varepsilon>0}$ to $(P_\alpha^\varepsilon)_{\varepsilon>0}$.

Let $(\varepsilon_n)_{n \geq 1}$, $\varepsilon_n \downarrow 0$, and $(\alpha_n)_{n \geq 1}$, $\alpha_n \in K$, $K \subset \mathcal{A}$ is a compact, be arbitrary sequences.

Proposition 2.1. *For each $h \in \mathcal{H}$ the sequence of measures $(P_{\alpha_n}^{\varepsilon_n,h})$ is contiguous to sequence of measures $(P_{\alpha_n}^{\varepsilon_n})$, i.e.*

$$(P_{\alpha_n}^{\varepsilon_n,h}) \triangleleft (P_{\alpha_n}^{\varepsilon_n}).$$

Proof. From the predictable criteria of contiguity (see, e.g., Jacod and Shiryaev [15]), follows that we have to verify the relation

$$\lim_{N \rightarrow \infty} \limsup_{n \rightarrow \infty} P_{\alpha_n}^{\varepsilon_n,h} \left\{ h_t^n \left(\frac{1}{2} \right) > N \right\} = 0, \quad (2.21)$$

where $h^n(\frac{1}{2}) = (h_s^n(\frac{1}{2}))_{0 \leq s \leq t}$ is the Hellinger process of order $\frac{1}{2}$.

By the definition of Hellinger process (see, e.g., Jacod and Shiryaev [15]) we have

$$h_t^n \left(\frac{1}{2} \right) = h_t^n \left(\frac{1}{2}, P_{\alpha_n}^{\varepsilon_n,h}, P_{\alpha_n}^{\varepsilon_n} \right) = \frac{1}{8} \int_0^t [h(s, \varepsilon_n X; \alpha_n)]^2 ds,$$

and since $h \in \mathcal{H}$, and hence is bounded, $h_t^n(\frac{1}{2})$ is bounded too, which provides (2.21). \square

Proposition 2.2. *For each estimate $(\alpha_t^{\varepsilon,\psi})_{\varepsilon>0}$ with $\psi \in \Psi_0$ and each alternative $(P_\alpha^{\varepsilon,h})_{\varepsilon>0} \in (\mathbf{P}_\alpha^{\varepsilon,\mathcal{H}})_{\varepsilon>0}$ the following relation holds true*

$$\mathcal{L} \left\{ \varepsilon^{-1}(\alpha_t^{\psi,\varepsilon} - \alpha) \mid P_\alpha^{\varepsilon,h} \right\} \xrightarrow{w} N \left([\gamma_0^\psi(\alpha)]^{-1} b(\psi, h; \alpha), V(\psi, \alpha) \right),$$

where

$$b(\psi, h; \alpha) := \int_0^t \psi(s, Y^0(\alpha); \alpha) h(s, Y^0(\alpha); \alpha) ds.$$

Proof. Proposition 2.1 together with (2.16) provides that uniformly in α on each compact

$$P_\alpha^{\varepsilon,h} - \lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} r_{\psi,\varepsilon}(\alpha) = 0,$$

and therefore we have to establish the limit distribution of random vector $[\gamma_0^\psi(\alpha)]^{-1} \varepsilon L_t^{\psi,\varepsilon}$ under the measures $(P_\alpha^{\varepsilon,h})_{\varepsilon>0}$.

By virtue of the Girsanov Theorem the process $L^{\psi,\varepsilon}(\alpha) = (L_s^{\psi,\varepsilon}(\alpha))_{0 \leq s \leq t}$ is a semimartingale with canonical decomposition

$$L_s^{\psi,\varepsilon}(\alpha) = \tilde{L}_s^{\psi,\varepsilon}(\alpha) + b_{\varepsilon,s}(\psi, h; \alpha), \quad 0 \leq s \leq t, \quad (2.22)$$

where $\tilde{L}^{\psi,\varepsilon}(\alpha) = (\tilde{L}_s^{\psi,\varepsilon}(\alpha))_{0 \leq s \leq t}$ is a $P_\alpha^{\varepsilon,h}$ -square integrable martingales defined as follows

$$\tilde{L}_s^{\psi,\varepsilon}(X; \alpha) := \int_0^s \psi_\varepsilon(u, X; \alpha) (dX_u - (a_\varepsilon(u, X; \alpha) + \varepsilon h_\varepsilon(u, X; \alpha)) du,$$

and

$$b_{\varepsilon,s}(\psi, h; \alpha) := \varepsilon \int_0^s \psi_\varepsilon(u, X; \alpha) h_\varepsilon(u, X; \alpha) du.$$

But $\langle \tilde{L}^{\psi,\varepsilon}(\alpha) \rangle_t = \Gamma_\varepsilon^\psi(\alpha)$, where $\Gamma_\varepsilon^\psi(\alpha)$ is defined by (2.8). On the other hand, from Proposition 2.1 and (2.10) it follows that

$$P_\alpha^{\varepsilon,h} - \lim_{\varepsilon \rightarrow 0} \langle \varepsilon \tilde{L}^{\psi,\varepsilon}(\alpha) \rangle_t = P_\alpha^{\varepsilon,h} - \lim_{\varepsilon \rightarrow 0} \varepsilon^2 \Gamma_\varepsilon^\psi(\alpha) = P_\alpha^\varepsilon - \lim_{\varepsilon \rightarrow 0} \varepsilon^2 \Gamma_\varepsilon^\psi(\alpha) = \Gamma_0^\psi(\alpha)$$

uniformly in α on each compact, and hence

$$\mathcal{L} \left\{ [\gamma_0^\psi(\alpha)]^{-1} \varepsilon \tilde{L}_t^{\psi,\varepsilon}(\alpha) \mid P_\alpha^{\varepsilon,h} \right\} \xrightarrow{w} N(0, V(\psi; \alpha)). \quad (2.23)$$

Finally, relation (2.23) together with (2.22) and relation

$$P_\theta^{\varepsilon,h} - \lim_{\varepsilon \rightarrow 0} \varepsilon b_{\varepsilon,t}(\psi, h; \alpha) = \int_0^t \psi(s, Y^0(\alpha); \alpha) h(s, Y^0(\alpha); \alpha) ds = b(\psi, h; \alpha),$$

provides the desirable results. \square

2.3. Optimization criteria. Construction of optimal B -robust estimates. In this subsection we state and solve an optimization problem, which results in construction of optimal B -robust estimate.

Initially, it should be stressed that the bias vector $\tilde{b}(\psi, h; \alpha) := [\gamma_0^\psi(\alpha)]^{-1} \times b(\psi, h; \alpha)$ can be viewed as the influence functional of the estimate $(\alpha_t^{\psi, \varepsilon})_{\varepsilon > 0}$ w.r.t. alternative $(P_\alpha^{\psi, h})_{\varepsilon > 0}$.

Indeed, the expansion (2.15) together with (2.22) and (2.23) allows to conclude that

$$\mathcal{L} \left\{ \varepsilon^{-1} (\alpha_t^{\psi, \varepsilon} - \alpha - \varepsilon^2 [\gamma_0^\psi(\alpha)]^{-1} b_\varepsilon(\psi, h; \alpha)) \mid P_\alpha^{\varepsilon, h} \right\} \xrightarrow{w} N(0, V(\psi, \alpha)),$$

and, hence, the expression

$$\alpha + \varepsilon^2 [\gamma_0^\psi(\alpha)]^{-1} b_\varepsilon(\psi, h; \alpha) - \alpha = \varepsilon^2 [\gamma_0^\psi(\alpha)]^{-1} b_\varepsilon(\psi, h; \alpha),$$

plays the role of bias on the “fixed step ε ” and it seems natural to interpret the limit

$$P_\alpha^{\varepsilon, h} - \lim_{\varepsilon \rightarrow 0} \frac{\alpha + \varepsilon^2 [\gamma_0^\psi(\alpha)]^{-1} b_\varepsilon(\psi, h; \alpha) - \alpha}{\varepsilon} = [\gamma_0^\psi(\alpha)]^{-1} b(\psi, h; \alpha),$$

as the influence functional.

For each estimate $(\alpha_t^{\psi, \varepsilon})_{\varepsilon > 0}$, $\psi \in \Psi_0$, define the risk functional w.r.t. alternative $(P_\alpha^{\varepsilon, h})_{\varepsilon > 0}$, $h \in \mathcal{H}$, as follows:

$$D(\psi, h; \alpha) = \lim_{a \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} E_\alpha^{\varepsilon, h} \left((\varepsilon^{-2} |\alpha_t^{\psi, \varepsilon} - \alpha|^2) \wedge a \right),$$

where $x \wedge a = \min(x, a)$, $a > 0$, $E_\alpha^{\varepsilon, h}$ is an expectation w.r.t. measure $P_\alpha^{\varepsilon, h}$.

Using Proposition 2.2 it is not hard to verify that

$$D(\psi, h; \alpha) = |\tilde{b}(\psi, h; \alpha)|^2 + \text{tr } V(\psi, \alpha),$$

where $\text{tr } A$ denotes the trace of matrix A .

By Proposition 2.2

$$\varepsilon^{-1} (\alpha_t^{\psi, \varepsilon} - \alpha) \xrightarrow{d} N(\tilde{b}(\psi, h; \alpha), V(\psi; \alpha)),$$

where \xrightarrow{d} denotes the convergence by distribution (by distribution $P_\alpha^{\varepsilon, h}$ in our case), $N(\tilde{b}, V)$ is a Gaussian random vector with mean \tilde{b} and covariation matrix V .

But if $\xi = (\xi_1, \dots, \xi_m)'$ is a Gaussian vector with parameters (μ, σ^2) , then

$$E|\xi|^2 = \sum_{i=1}^m E\xi_i^2 = \sum_{i=1}^m (E\xi_i)^2 + \sum_{i=1}^m D\xi_i = |\mu|^2 + \text{tr } \sigma^2,$$

as it was required.

Connect with each $\psi \in \Psi_0$ the function $\tilde{\psi}$ as follows

$$\tilde{\psi}(s, x; \alpha) = [\gamma_0^\psi(\alpha)]^{-1} \psi(s, x; \alpha), \quad 0 \leq s \leq t, \quad x \in C_t, \quad \alpha \in \mathcal{A}.$$

Then $\tilde{\psi} \in \Psi_0$ and

$$\gamma_0^{\tilde{\psi}}(\alpha) = Id,$$

where Id is an unit matrix,

$$V(\psi; \alpha) = V(\tilde{\psi}; \alpha) = \Gamma_0^{\tilde{\psi}}(\alpha), \quad \tilde{b}(\psi, h; \alpha) = \tilde{b}(\tilde{\psi}, h; \alpha) = b(\tilde{\psi}, h; \alpha).$$

Therefore

$$D(\psi, h; \alpha) = D(\tilde{\psi}, h; \alpha) = |b(\tilde{\psi}, h; \alpha)|^2 + \text{tr} \Gamma_0^{\tilde{\psi}}(\alpha). \quad (2.24)$$

Denote \mathcal{H}_r , a set of functions $h \in \mathcal{H}$ such that for each $\alpha \in \mathcal{A}$

$$\int_0^t |h(s, Y^0(\alpha); \alpha)| ds \leq r,$$

where $r, r > 0$, is a constant.

Since, for each $r > 0$,

$$\sup_{h \in \mathcal{H}_r} |b(\tilde{\psi}, h; \alpha)| \leq \text{const}(r) \sup_{0 \leq s \leq t} |\tilde{\psi}(s, Y^0(\alpha); \alpha)|,$$

where constant depends on r , we call the function $\tilde{\psi}$ an influence function of estimate $(\alpha_t^{\psi, \varepsilon})_{\varepsilon > 0}$ and a quantity

$$\gamma_\psi^*(\alpha) = \sup_{0 \leq s \leq t} |\tilde{\psi}(s, Y^0(\alpha); \alpha)|$$

is named as the (unstandardized) gross error sensitivity at point α of estimate $(\alpha_t^{\psi, \varepsilon})_{\varepsilon > 0}$.

Define

$$\Psi_{0,c} = \left\{ \psi \in \Psi_0 : \int_0^t \psi(s, Y^0(\alpha); \alpha) [\dot{a}(s, Y^0(\alpha); \alpha)]' ds = Id, \quad (2.25) \right.$$

$$\left. \gamma_\psi^*(\alpha) \leq c \right\}, \quad (2.26)$$

where $c \in [0, \infty)$ is a generic constant.

Take into account the expression (2.24) for the risk functional we come to the following optimization problem, known in robust estimation theory as Hampel's optimization problem: minimize the trace of the asymptotic covariance matrix of estimate $(\alpha_t^{\psi, \varepsilon})_{\varepsilon > 0}$ over the class $\Psi_{0,c}$, that is

$$\text{minimize} \quad \int_0^t \psi(s, Y^0(\alpha); \alpha) [\psi(s, Y^0(\alpha); \alpha)]' ds \quad (2.27)$$

under the side conditions (2.25) and (2.26).

Define the Huber function $h_c(z)$, $z \in R^m$, $c > 0$, as follows

$$h_c(z) := z \min \left(1, \frac{c}{|z|} \right).$$

For arbitrary nondegenerate matrix A denote $\psi_c^A = h_c(A\dot{a})$.

Theorem 2.1. *Assume that for given constant c there exists a nondegenerate $m \times m$ -matrix $A_c^*(\alpha)$, which solves the equation (w.r.t. matrix A)*

$$\int_0^t \psi_c^A(s, Y^0(\alpha); \alpha) [\dot{a}(s, Y^0(\alpha); \alpha)]' ds = Id. \quad (2.28)$$

Then the function $\psi_c^{A_c^(\alpha)} = h_c(A_c^*(\alpha)\dot{a})$ solves the optimization problem (2.27).*

Proof. We follow Hampel et al. [12].

Let A be an arbitrary $m \times m$ -matrix.

Since for each $\psi \in \Psi_{0,c}$, $\int \psi(\dot{a})' = Id$, $\int \dot{a}[\dot{a}]' = I^0(\alpha)$ (see (2.7)), then

$$\int (\psi - A\dot{a})(\psi - A\dot{a})' = \int \psi\psi' - A - A' + AI^0(\alpha)A'$$

(here and below we use simple evident notation for integrals).

Therefore since the trace is an additive functional instead of minimizing of $\text{tr} \int \psi\psi'$ we can minimize

$$\text{tr} \int (\psi - A\dot{a})(\psi - A\dot{a})' = \int |\psi - A\dot{a}|^2.$$

Note that for each z

$$\arg \min_{|y| \leq c} |z - y|^2 = h_c(z).$$

Indeed, it is evident that minimizing y has the form $y = \beta z$, where β , $0 \leq \beta \leq 1$, is constant. Then

$$\min_{|y| \leq c} |z - y|^2 = \min_{\beta \leq \frac{c}{|z|}} (1 - \beta)^2 |z|^2.$$

Thus we have to find

$$\arg \min_{\beta \leq \frac{c}{|z|}} (1 - \beta)^2 = \min \left(1, \frac{c}{|z|} \right).$$

But last relation is trivially satisfied. Hence the minimizing $y^* = z \min(1, \frac{c}{|z|})$ and

$$\arg \min_{|\psi| \leq c} |\psi - A\dot{a}|^2 = h_c(A\dot{a}).$$

From the other side,

$$|h_c(z)|^2 = |z|^2 I_{\{|z| \leq c\}} + \frac{|z|^2}{|z|^2} c^2 I_{\{|z| \geq c\}} \leq c^2.$$

Hence

$$|h_c(z)| \leq c \quad \text{for all } z$$

and therefore $h_c(A\dot{a})$ satisfies the condition (2.26) for each A .

Now it is evident that a function $h_c(A\dot{a})$ minimizes the expression under integral sign, and hence the integral itself over all functions $\psi \in \Psi_0$ satisfying (2.26).

At the same time the condition (2.25), generally speaking, can be violated. But, since a matrix A is arbitrary, we can choose $A = A_c^*(\alpha)$ from (2.28) which, of course, guarantees the validity of (2.25) for $\psi_c^* = \psi_c^{A_c^*(\alpha)}$. \square

As we have seen the resulting optimal influence functions ψ_c^* is defined along the process $Y^0(\alpha) = (Y_s^0(\alpha))_{0 \leq s \leq t}$, which is a solution of equation (2.2).

But for constructing optimal estimate we need a function $\psi_c^*(s, x; \alpha)$ defined on whole space $[0, t] \times C_t \times \mathcal{A}$.

For this purpose define $\psi_c^*(s, x; \alpha)$ as follows;

$$\psi_c^*(s, x; \alpha) = \psi_c^{A_c^*(\alpha)}(s, x; \alpha) = h_\varepsilon(A_c^*(\alpha)\dot{a}(s, x; \alpha)), \quad (2.29)$$

and as usual $\psi_{c,\varepsilon}^*(s, x; \alpha) = \frac{1}{\varepsilon}\psi_c^*(s, \varepsilon x; \alpha)$, $0 \leq s \leq t$, $x \in C_t$, $\alpha \in \mathcal{A}$.

Definition 2.2. We say that $\psi_c^*(s, x; \alpha)$, $0 \leq s \leq t$, $x \in C_t$, $\alpha \in \mathcal{A}$, is an influence function of optimal B -robust estimate $(\alpha_t^{*,\varepsilon})_{\varepsilon>0} = (\alpha_t^{\psi_c^*,\varepsilon})_{\varepsilon>0}$ over the class of CULAN estimates $(\alpha_t^{\psi,\varepsilon})_{\varepsilon>0}$, $\psi \in \Psi_{0,c}$, if the matrix $A^*(\alpha)$ is differentiable in α .

From (2.9), (2.11), (2.28) and (2.29) it directly follows that

$$\gamma_0^{\psi_c^*}(\alpha) = P_\alpha^\varepsilon - \lim_{\varepsilon \rightarrow 0} \varepsilon^2 \gamma_\varepsilon^{\psi_c^*}(\alpha) = \int_0^t \psi_c^*(s, Y^0(\alpha); \alpha) (\dot{a}(s, Y^0(\alpha); \alpha))' ds = Id.$$

Besides, for each alternative $(P_\alpha^{\varepsilon,h})_{\varepsilon>0}$, $h \in \mathcal{H}$, according to the Proposition 2.2 we have

$$\mathcal{L} \{ \varepsilon^{-1}(\alpha_t^{*,\varepsilon} - \alpha) \mid P_\alpha^{\varepsilon,h} \} \xrightarrow{w} N(b(\psi_c^*, h; \alpha), V(\psi_c^*; \alpha)) \quad \text{as } \varepsilon \rightarrow 0,$$

where

$$b(\psi_c^*, h; \alpha) = \int_0^t \psi_c^*(s, Y^0(\alpha); \alpha) h(s, Y^0(\alpha); \alpha) ds,$$

and $V(\psi_c^*; \alpha) = \Gamma_0^{\psi_c^*}(\alpha)$.

Hence, the risk functional for estimate $(\alpha_t^{*,\varepsilon})_{\varepsilon>0}$ is

$$D(\psi_c^*, h; \alpha) = |b(\psi_c^*, h; \alpha)|^2 + \text{tr} \Gamma_0^{\psi_c^*}, \quad h \in \mathcal{H},$$

and the (unstandardized) gross error sensitivity of $(\alpha_t^{*,\varepsilon})_{\varepsilon>0}$ is

$$\gamma_{\psi_c^*}(\alpha) = \sup_{0 \leq s \leq t} |\psi_c^*(s, Y^0(\alpha); \alpha)| \leq c.$$

From above reasons, we may conclude that $(\alpha_t^{*,\varepsilon})_{\varepsilon>0}$ is the optimal B -robust estimate over the class of estimates $(\alpha_t^{\psi,\varepsilon})_{\varepsilon>0}$, $\psi \in \Psi_{0,c}$ in the following sense: the trace of asymptotic covariance matrix of $(\alpha_t^{*,\varepsilon})_{\varepsilon>0}$ is minimal among all estimates $(\alpha_t^{\psi,\varepsilon})_{\varepsilon>0}$ with bounded by constant gross error sensitivity, that is

$$\Gamma_0^{\psi_c^*}(\alpha) \leq \Gamma_0^\psi(\alpha) \quad \text{for all } \psi \in \Psi_{0,c}. \quad \square$$

Note that for each estimate $(\alpha_t^{\psi, \varepsilon})_{\varepsilon > 0}$ and alternatives $(P_\alpha^{\varepsilon, h})_{\varepsilon > 0}$, $h \in \mathcal{H}$, the influence functional is bounded by $\text{const}(r) \cdot c$. Indeed, we have for $\psi \in \Psi_{0, c}$,

$$\sup_{h \in \mathcal{H}_r} |b(\psi, h; \alpha)| \leq \text{const}(r) \cdot c := C(r, c),$$

and since from (2.24)

$$\inf_{\psi \in \Psi_{0, c}} \sup_{h \in \mathcal{H}_r} D(\psi, h; \alpha) \leq C^2(r, c) + \text{tr} \Gamma_0^{\psi_c^*}(\alpha),$$

we can choose “optimal level” of truncation, minimizing the expression

$$C^2(r, c) + \text{tr} \Gamma_0^{\psi_c^*}(\alpha)$$

over all constants c , for which the equation (2.28) has a solution $A_c^*(\alpha)$. This can be done using the numerical methods.

For the problem of existence and uniqueness of solution of equation (2.28) we address to Rieder [34].

In the case of one-dimensional parameter α (i.e. $m = 1$) the optimal level c^* of truncation is given as an unique solution of the following equation (see Lazrieva and Toronjadze [20], [21])

$$r^2 c^2 = \int_0^t [\dot{a}(s, Y^0(\alpha); \alpha)]_{-c}^c \dot{a}(s, Y^0(\alpha); \alpha) ds - \int_0^t ([\dot{a}(s, Y^0(\alpha); \alpha)]_{-c}^c)^2 ds,$$

where $[x]_a^b = (x \wedge b) \vee a$ and the resulting function

$$\psi^*(s, x; \alpha) = [\dot{a}(s, x; \alpha)]_{-c^*}^{c^*}, \quad 0 \leq s \leq t, \quad x \in C_t,$$

is (Ψ_0, \mathcal{H}_r) optimal in the following minimax sense:

$$\sup_{h \in \mathcal{H}_r} D(\psi^*, h; \alpha) = \inf_{\psi \in \Psi} \sup_{h \in \mathcal{H}_r} D(\psi, h; \alpha).$$

3. OPTIMAL MEAN-VARIANCE ROBUST HEDGING

3.1. A financial market model. Let $(\Omega, \mathcal{F}, F = (\mathcal{F}_t)_{0 \leq t \leq T}, P)$ be a filtered probability space with filtration F satisfying the usual conditions, where $T \in (0, \infty]$ is a fixed time horizon. Assume that \mathcal{F}_0 is a trivial and $\mathcal{F}_T = \mathcal{F}$.

There exist $d + 1$, $d \geq 1$ primitive assets: one bond, whose price process is assumed to be 1 at all times and d risky assets (stocks), whose R^d -valued price process $X = (X_t)_{0 \leq t \leq T}$ is a continuous semimartingale given by the relation:

$$dX_t = \text{diag}(X_t) dR_t, \quad X_0 > 0, \quad (3.1)$$

where $\text{diag}(X)$ denotes the diagonal $d \times d$ -matrix with diagonal elements X^1, \dots, X^d , and the yield process $R = (R_t)_{0 \leq t \leq T}$ is a R^d -valued continuous semimartingale satisfying the stricture condition (SC). That is (see Schweizer [37])

$$dR_t = d\langle \widetilde{M} \rangle_t \lambda_t + d\widetilde{M}_t, \quad R_0 = 0, \quad (3.2)$$

where $\widetilde{M} = (\widetilde{M}_t)_{0 \leq t \leq T}$ is a R^d -valued continuous martingale, $\widetilde{M} \in \mathcal{M}_{0,\text{loc}}^2(P)$, $\lambda = (\lambda_t)_{0 \leq t \leq T}$ is a F -predictable R^d -valued process, and the mean-variance tradeoff (MVT) process $\widetilde{\mathcal{K}} = (\widetilde{\mathcal{K}}_t)_{0 \leq t \leq T}$ of process R

$$\widetilde{\mathcal{K}}_t := \int_0^t \lambda'_s d\langle \widetilde{M} \rangle_s \lambda_s = \langle \lambda' \cdot \widetilde{M} \rangle_t < \infty \quad P\text{-a.s.}, \quad t \in [0, T]. \quad (3.3)$$

Remark 3.1. Remember that all vectors are assumed to be column vectors.

Suppose that the martingale \widetilde{M} has the form

$$\widetilde{M} = \sigma \cdot M, \quad (3.4)$$

where $M = (M_t)_{0 \leq t \leq T}$ is a R^d -valued continuous martingale, $M \in \mathcal{M}_{0,\text{loc}}^2(P)$ with $d\langle M^i, M^j \rangle_t = I_{ij}^{d \times d} dC_t$, $I^{d \times d}$ is the identity matrix, $C = (C_t)_{0 \leq t \leq T}$ is a continuous increasing bounded process with $C_0 = 0$.

Further, let $\sigma = (\sigma_t)_{0 \leq t \leq T}$ is a $d \times d$ -matrix valued, F -predictable process with $\text{rank}(\sigma_t) = d$ for any t , P -a.s., the process $(\sigma_t^{-1})_{0 \leq t \leq T}$ is locally bounded, and

$$\int_0^T \sigma_t d\langle M \rangle_t \sigma'_t < \infty \quad P\text{-a.s.} \quad (3.5)$$

Assume now that the following condition be satisfied:

There exist fixed R^d -valued, F -predictable process $k = (k_t)_{0 \leq t \leq T}$ such that

$$\lambda = \lambda(\sigma) = (\sigma')^{-1} k. \quad (3.6)$$

In the case from (3.2) we get

$$\begin{aligned} dR_t &= d\langle \widetilde{M} \rangle_t \lambda_t + d\widetilde{M}_t = \sigma_t d\langle M \rangle_t \sigma'_t (\sigma'_t)^{-1} k_t + \sigma_t dM_t \\ &= \sigma_t (d\langle M \rangle_t k_t + dM_t), \end{aligned} \quad (3.7)$$

and

$$\begin{aligned} \widetilde{\mathcal{K}}_t &= \int_0^t \lambda'_s d\langle \widetilde{M} \rangle_s \lambda_s = \int_0^t k'_t ((\sigma'_t)^{-1})' \sigma_t d\langle \widetilde{M} \rangle_t \sigma'_t (\sigma'_t)^{-1} k_t \\ &= \int_0^t k'_t d\langle M \rangle_t k_t = \langle k \cdot M \rangle_t := \mathcal{K}_t. \end{aligned}$$

From (3.3) we have

$$\mathcal{K}_t < \infty \quad P\text{-a.s.} \quad \text{for all } t \in [0, T]. \quad (3.8)$$

Thus, if we introduce the process $M^0 = (M_t^0)_{0 \leq t \leq T}$ by the relation

$$dM_t^0 = d\langle M \rangle_t k_t + dM_t, \quad M_0^0 = 0, \quad (3.9)$$

then the MVT process $\mathcal{K} = (\mathcal{K}_t)_{0 \leq t \leq T}$ of R^d -valued semimartingale M^0 is finite, and hence M^0 satisfies SC.

Finally, the scheme (3.1), (3.2), (3.4), (3.6) and (3.9) can be rewritten in the following form

$$\begin{aligned} dX_t &= \text{diag}(X_t) dR_t, \quad X_0 > 0, \\ dR_t &= \sigma_t dM_t^0, \quad R_0 = 0, \\ dM_t^0 &= d\langle M \rangle_t k_t + dM_t, \quad M_0 = 0, \end{aligned} \tag{3.10}$$

where σ and k satisfy (3.5) and (3.8), respectively.

This is our financial market model.

3.2. Characterization of variance-optimal ELMM (equivalent local martingale measure). A key role in mean-variance hedging plays variance-optimal ELMM (see, e.g., RSch [33], GLP [11]). Here we collect some facts characterizing this measure.

We start with remark that the sets ELMMs for processes X , R and M^0 form (3.10) coincide. Hence we can and will consider the simplest process M^0 .

Introduce the notation

$$\mathcal{M}_2^e := \left\{ Q \sim P : \frac{dQ}{dP} \in L^2(P), \quad M^0 \text{ is a } Q\text{-local martingale} \right\},$$

and suppose that

$$(c.1) \quad \mathcal{M}_2^e \neq \emptyset.$$

The solution \tilde{P} of the optimization problem

$$E\mathcal{E}_T^2(\mathcal{M}^Q) \rightarrow \inf_{Q \in \mathcal{M}_2^e}$$

is called variance-optimal ELMM.

Here

$$\frac{dQ}{dP} \Big|_{\mathcal{F}_T} = \mathcal{E}_T(M^Q),$$

and $(\mathcal{E}_t(M^Q))_{0 \leq t \leq T}$ is the Dolean exponential of martingale M^Q .

It is well-known (see, e.g., Schweizer [37], [38]) that under condition (c.1) variance-optimal ELMM \tilde{P} exist.

Denote

$$\tilde{z}_T := \frac{d\tilde{P}}{dP} \Big|_{\mathcal{F}_T},$$

and introduce RCLL process $\tilde{z} = (\tilde{z}_t)_{0 \leq t \leq T}$ by the relation

$$\tilde{z}_t = E^{\tilde{P}}(\tilde{z}_T / \mathcal{F}_t), \quad 0 \leq t \leq T.$$

Then, by Schweizer [37], [38]

$$\tilde{z}_T = \tilde{z}_0 + \int_0^T \zeta'_t dM_t^0, \tag{3.11}$$

where $\zeta = (\zeta_t)_{0 \leq t \leq T}$ is the R^d -valued F -predictable process with

$$\int_0^T \zeta'_t d\langle M \rangle_t \zeta_t < \infty,$$

and the process $(\int_0^t \zeta'_s dM_s^0)_{0 \leq t \leq T}$ is a \tilde{P} -martingale.

Relation (3.11) easily implies that the process \tilde{z} is actually continuous.

Suppose, in addition to (c.1), that the following condition is satisfied:

(c.*) all P -local martingales are continuous.

This technical assumption is satisfied in stochastic volatility models, where $F = F^w$ is the natural filtration generated by the Wiener process.

It shown in Mania and Tevzadze [34], Mania et al. [25] that under conditions (c.1) and (c*) density \tilde{z}_T of variance optimal ELMM is uniquely characterized by the relation

$$\tilde{z}_T = \frac{\mathcal{E}_T((\varphi - k)' \cdot M^0)}{E\mathcal{E}_T((\varphi - k)' \cdot M^0)}, \quad (3.12)$$

where φ together with the pair (L, c) is the unique solution of the following equation

$$\frac{\mathcal{E}_T((\varphi - 2k)' \cdot M)}{\mathcal{E}_T(L)} = c\mathcal{E}_T^2(-k' \cdot M), \quad (3.13)$$

where $L \in M_{0,\text{loc}}^2(P)$, $\langle L, M \rangle = 0$, c is a constant.

Moreover, the process $\zeta = (\zeta_t)_{0 \leq t \leq T}$ from (3.11) has the form

$$\zeta_t = (\varphi_t - k_t)\mathcal{E}_t((\varphi - k)' \cdot M^0). \quad (3.14)$$

Here $\varphi = (\varphi_t)_{0 \leq t \leq T}$ is a R^d -valued, F -predictable process with

$$\int_0^T \varphi'_t d\langle M \rangle_t \varphi_t < \infty.$$

Let τ be F -stopping time.

Denote $\langle k' \cdot M \rangle_{T\tau} = \langle k' \cdot M \rangle_T - \langle k' \cdot M \rangle_\tau$.

Proposition 3.1 (see also Biagini et al. [3], LLaurent and Pham [18]).

1. Equation (3.13) is equivalent to equation

$$\frac{\mathcal{E}_T(\varphi' \cdot M^*)}{\mathcal{E}_T(L)} = ce^{\langle k' \cdot M \rangle_T}, \quad (3.15)$$

where the R^d -valued process $M^* = (M_t^*)_{0 \leq t \leq T}$ is given by the relation

$$dM_t^* = 2d\langle M \rangle_t k_t + dM_t, \quad M_0^* = 0.$$

2. a) If there exists the martingale $m = (m_t)_{0 \leq t \leq T}$, $m \in \mathcal{M}_{0,\text{loc}}^2(P)$ such that

$$e^{-\langle k' \cdot M \rangle_T} = c + m_T, \quad \langle m, M \rangle = 0, \quad (3.16)$$

then $\varphi \equiv 0$ and $L_T = \int_0^T \frac{1}{c+m} dm_s$ solve the equation (3.15).

In this case

$$\tilde{z}_T = \frac{\mathcal{E}_T(-k' \cdot M^0)}{E\mathcal{E}_T(-k' \cdot M^0)}, \quad (3.17)$$

process $\zeta = (\zeta_t)_{0 \leq t \leq T}$ from (3.11) is equal to

$$\zeta_t = -k_t \mathcal{E}_t(-k' \cdot M^0),$$

and

$$E \left[\left(\frac{\tilde{z}_T}{\tilde{z}_\tau} \right)^2 / \mathcal{F}_\tau \right] = \frac{1}{E(e^{-\langle k' \cdot M \rangle_{T\tau}} / \mathcal{F}_\tau)}.$$

b) If there exist R^d -valued F -predictable process $\ell = (\ell_t)_{0 \leq t \leq T}$, $\int_0^T \ell'_t d\langle M \rangle \ell_t < \infty$ and

$$e^{\langle k' \cdot M \rangle_T} = c + \int_0^T \ell'_t dM_t^*,$$

then $L \equiv 0$ and $\varphi_t = \frac{\ell_t}{c + \int_0^t \ell'_s dM_s^*}$ solve the equation (3.15).

In this case

$$\tilde{z}_T = \mathcal{E}_T(-k' \cdot M) \quad (:= \hat{z}_T, \text{ the density of minimal martingale measure } \hat{P}),$$

and

$$E \left(\left(\frac{\tilde{z}_T}{\tilde{z}_\tau} \right)^2 / \mathcal{F}_\tau \right) = E^{P^*}(e^{\langle k' \cdot M \rangle_{T\tau}} / \mathcal{F}_\tau),$$

where $dP^* = \mathcal{E}_T(-2k' \cdot M)dP$.

Proof. 1. By the Yor formula

$$\begin{aligned} \mathcal{E}_T(\varphi - 2k)' \cdot M &= \mathcal{E}_T(\psi' \cdot M - 2k' \cdot M) \\ &= \mathcal{E}_T \left(\varphi' \cdot \left(M + 2 \int_0^\cdot d\langle M \rangle_t k_t \right) - 2 \int_0^\cdot \psi'_t d\langle M \rangle_t k_t - 2k' \cdot M \right) \\ &= \mathcal{E}_T(\varphi' \cdot M^*) \mathcal{E}_T(-2k' \cdot M), \end{aligned}$$

and

$$\mathcal{E}_T^2(-k' \cdot M) = \mathcal{E}_T(-2k' \cdot M) e^{\langle k' \cdot M \rangle_T}.$$

Assertion follows.

2. a) Note at first that $\langle L, M \rangle = 0$. Further, by the formula we can write

$$\ln(c + m_t) - \ln c = \int_0^t \frac{1}{c + m_s} dm_s - \frac{1}{2} \int_0^t \frac{1}{(c + m_s)^2} d\langle m \rangle_s.$$

Hence

$$e^{\ln(c+m_T) - \ln c} = \mathcal{E}_T(L),$$

and thus

$$\mathcal{E}_T(L) = \frac{c + m_T}{c} = \frac{e^{-\langle k' \cdot M \rangle_T}}{c}.$$

Finally, by the Bayes rule and the Girsanov Theorem

$$\begin{aligned} E \left(\left(\frac{\tilde{z}_T}{\tilde{z}_\tau} \right)^2 / \mathcal{F}_\tau \right) &= \frac{E(\mathcal{E}_T(-2k' \cdot M) e^{-\langle k' \cdot M \rangle_T} / \mathcal{F}_\tau)}{E^2(\mathcal{E}_T(-k' \cdot M) e^{-\langle k' \cdot M \rangle_T} / \mathcal{F}_\tau)} \\ &= \frac{E^*(c + m_T / \mathcal{F}_\tau) \mathcal{E}_T^2(-k' \cdot M)}{(\widehat{E}(c + m_\tau / \mathcal{F}_\tau))^2 \mathcal{E}_T^2(-2k' \cdot M)} = \frac{c + m_\tau}{(c + m_\tau)^2} \cdot e^{\langle k' \cdot M \rangle_\tau} \\ &= \frac{1}{E(e^{\langle k' \cdot M \rangle_{T\tau}} / \mathcal{F}_\tau)}. \end{aligned}$$

The proof of case 2 b) is quite analogous.

Proposition is proved. \square

3.3. Misspecified asset price model and robust hedging. Denote by $\text{Ball}_L(0, r)$, $r \in [0, \infty)$ the closed r -radius ball in the space $L = L_\infty(dt \times dP)$, with the center at the origin, and let

$$\mathcal{H} := \{h = \{h_{ij}\}, i, j = 1, d : h \text{ is } F\text{-predictable } d \times d\text{-matrix valued process, } \text{rank}(h) = d, h_{ij} \in \text{Ball}_L(0, r), r \in [0, \infty)\}. \quad (3.18)$$

Class \mathcal{H} is called the class of alternatives.

Fix the value of small parameter $\delta > 0$, as well as $d \times d$ -matrix valued, F -predictable process $\sigma^0 = (\sigma_t^0)_{0 \leq t \leq T} = (\{\sigma_{ij,t}^0\}, 1 \leq i, j \leq d)_t$ such that $|\sigma_{ij,t}^0| \leq \text{const}$, $\forall i, j, t$, the matrix $(\sigma^0)^2 = \sigma^0(\sigma^0)'$ is uniformly elliptic, i.e. for each vector $v_t = (v_t^1, \dots, v_t^d)$ with probability 1

$$\sum_{i,j=1}^d (\sigma^0)_{ij,t}^2 v_t^i v_t^j \geq c \sum_{i=1}^d |v_t^i|^2, \quad c > 0, \quad 0 \leq t \leq T, \quad (3.19)$$

and denote

$$A_\delta = \{\sigma : \sigma = \sigma^0 + \delta h, \quad h \in \mathcal{H}\}. \quad (3.20)$$

Proposition 3.2. *Every σ from the class A_δ for sufficiently small δ is F -predictable $d \times d$ -valued process with bounded elements and the matrix $\sigma^2 = \sigma \sigma'$ is uniformly elliptic.*

Proof. The process σ is F -predictable as linear combination of F -predictable processes. Further,

$$|\sigma_{ij,t}| = |\sigma_{ij,t}^0 + \delta h_{ij,t}| \leq \text{const} + \delta r, \quad 0 < \delta \ll 1.$$

From (3.19) and (3.20) for each vector $v_t = (v_t^1, \dots, v_t^d)$ we have

$$\sum_{i,j=1}^d (\sigma^2)_{ij,t} v_t^i v_t^j = \sum_{i,j=1}^d (\sigma^0 + \delta h)(\sigma^0 + \delta h)'_{ij,t} v_t^i v_t^j$$

$$\begin{aligned}
&= \sum_{i,j=1}^d (\sigma^0 (\sigma^0)')_{ij,t} \nu_t^i \nu_t^j + \delta \sum_{i,j=1}^d (\sigma^0 h')_{ij,t} \nu_t^i \nu_t^j \\
&\quad + \delta \sum_{i,j=1}^d (h (\sigma^0)')_{ij,t} \nu_t^i \nu_t^j + \delta^2 \sum_{i,j=1}^d (hh')_{ij,t} \nu_t^i \nu_t^j. \tag{3.21}
\end{aligned}$$

Note now that the elements of matrices σ^0 and h are bounded. Hence choosing δ sufficiently small we get

$$\max(|\delta|(\sigma^0 h')_{ij,t}|, |\delta|(h(\sigma^0)')_{ij,t}|, \delta^2|(hh')_{ij,t}|) \leq \frac{\varepsilon}{3}.$$

Therefore from (3.19) and (3.21) we get

$$\sum_{i,j=1}^d \sigma_{ij,t}^2 \nu_t^i \nu_t^j \geq (c - \text{const} \cdot \varepsilon) \sum_{i,j=1}^d |\nu_t^i|^2 \quad \text{for each } \varepsilon > 0.$$

Proposition is proved. \square

Consider the set of processes $\{R^\sigma(\text{or } X^\sigma), \sigma \in A_\delta\}$, which represents the misspecified of asset price model.

Define the class of admissible trading strategies $\Theta = \Theta(\sigma^0)$.

Proposition 3.3. *For each R^d -valued F -predictable process $\theta = (\theta_t)_{0 \leq t \leq T}$ and for each $\sigma \in A_\delta$, $\delta > 0$,*

$$\begin{aligned}
aE \int_0^T |\theta_t|^2 dC_t &\leq E \int_0^T \theta'_t \sigma_t d\langle M \rangle_t \sigma'_t \theta_t \\
&= E \int_0^T \theta'_t \sigma_t \sigma'_t \theta_t dC_t \leq AE \int_0^T |\theta_t|^2 dC_t,
\end{aligned}$$

where the constants a, A are such that $0 < a \leq A < \infty$, and the parameter $\delta > 0$ is sufficiently small.

Proof. Remember that $d\langle M \rangle_t = d\langle M^i, M^j \rangle_t = I_{ij}^{d \times d} dC_t$. Hence

$$E \int_0^T \theta'_t \sigma_t d\langle M \rangle_t \sigma'_t \theta_t = E \int_0^T \theta'_t \sigma_t \sigma'_t \theta_t dC_t.$$

Further, since $\sigma = \sigma^0 + \delta h$ and elements of matrices σ^0 and h are bounded, then the same is true for the elements of matrix σ with $0 \leq \delta \leq \text{const}$. Thus using the inequality $ab \leq 2(a^2 + b^2)$ we get

$$E \int_0^T \theta'_t \sigma_t \sigma'_t \theta_t dC_t \leq AE \int_0^T |\theta_t|^2 dC_t.$$

On the other hand, by Proposition 3.2 the matrix $\sigma^2 = \sigma \sigma'$ is uniformly elliptic for sufficiently small δ , which yields the first inequality. \square

Definition 3.1. The class $\Theta = \Theta(\sigma^0)$ is a class of R^d -valued F -predictable processes $\theta = (\theta_t)_{0 \leq t \leq T}$ such that

$$E \int_0^T |\theta_t|^2 dC_t < \infty. \quad (3.22)$$

Let $\theta \in \Theta$ be the dollar amount (rather than the number of shares) invested in the stock X^σ , $\sigma \in A_\delta$. Then for each $\sigma \in A_\delta$ the trading gains induced by the self-financing portfolio strategy associated to θ has the form

$$G_t(\sigma, \theta) = \int_0^t \theta'_s dR_s^\sigma, \quad 0 \leq t \leq T, \quad (3.23)$$

where $R^d = (R_t^d)_{0 \leq t \leq T}$ is the yield process given by (3.10).

Introduce the condition:

(c.2) There exists ELMM \overline{Q} such that the density process $z = z^{\overline{Q}}$ satisfies the reverse Hölder inequality $R_2(P)$, see definition in RSch [33].

It is well-known that under the conditions (c.1) and (c.2) the density process $\tilde{z} = (\tilde{z}_t)_{t \leq T}$ of the variance-optimal ELMM satisfies $R_2(P)$ as well, see Dolean et al. [8].

Now under the conditions (c.1) and (c.2) the r.v. $G_T(\sigma, \theta) \in L^2(P)$, $\forall \sigma \in A_\delta$, and the space $G_T(\sigma, \Theta)$ is closed in $L^2(P)$, $\forall \sigma \in A_\delta$ (see, e.g., Theorem 2 of RSch [33]).

A contingent claim is an \mathcal{F}_T -measurable square-integrable r.v. H , which models the payoff from a financial product at the maturity date T .

The problem we are interested in is to find the robust hedging strategy for a contingent claim H in the above described incomplete financial market model with misspecified asset price process X^σ , $\sigma \in A_\delta$, using mean-variance approach.

For each $\sigma \in A_\delta$, the total loss of a hedger, who starts with the initial capital x , uses the strategy θ , believes that the stock price process follows X^σ , and has to pay a random amount H at the date T , is $H - x - G_T(\sigma, \theta)$.

Denote

$$\mathcal{J}(\sigma, \theta) := E(H - x - G_T(\sigma, \theta))^2. \quad (3.24)$$

One setting of the robust mean-variance hedging problem consist in solving the optimization problem

$$\text{minimize } \sup_{\sigma \in A_\delta} \mathcal{J}(\sigma, \theta) \text{ over all strategies } \theta \in \Theta. \quad (3.25)$$

We “slightly” change this problem using the approach developed in Toronjadze [41] which based on the following approximation

$$\begin{aligned} \sup_{\sigma \in A_\delta} \mathcal{J}(\sigma, \theta) &= \exp \left\{ \sup_{h \in \mathcal{H}} \ln \mathcal{J}(\sigma^0 + \delta h, \theta) \right\} \\ &\simeq \exp \left\{ \sup_{h \in \mathcal{H}} \left[\ln \mathcal{J}(\sigma^0, \theta) + \delta \frac{D\mathcal{J}(\sigma^0, h, \theta)}{\mathcal{J}(\sigma^0, \theta)} \right] \right\} \end{aligned}$$

$$= \mathcal{J}(\sigma^0, \theta) \exp \left\{ \delta \sup_{h \in \mathcal{H}} \frac{D\mathcal{J}(\sigma^0, h, \theta)}{\mathcal{J}(\sigma^0, \theta)} \right\},$$

where

$$D\mathcal{J}(\sigma^0, h, \theta) := \frac{d}{d\delta} \mathcal{J}(\sigma^0 + \delta h, \theta)|_{\delta=0} = \lim_{\delta \rightarrow 0} \frac{\mathcal{J}(\sigma^0 + \delta h, \theta) - \mathcal{J}(\sigma^0, \theta)}{\delta},$$

is the Gateaux differential of the functional \mathcal{J} at the point σ^0 in the direction h .

Approximate (in leading order δ) the optimization problem (3.25) by the problem

$$\begin{aligned} & \text{minimize } \mathcal{J}(\sigma^0, \theta) \exp \left\{ \delta \sup_{h \in \mathcal{H}} \frac{D\mathcal{J}(\sigma^0, h, \theta)}{\mathcal{J}(\sigma^0, \theta)} \right\} \\ & \text{over all strategies } \theta \in \Theta. \end{aligned} \quad (3.26)$$

Note that each solution θ^* of the problem (3.26) minimizes $\mathcal{J}(\sigma^0, \theta)$ under the constraint

$$\sup_{h \in \mathcal{H}} \frac{D\mathcal{J}(\sigma^0, h, \theta)}{\mathcal{J}(\sigma^0, \theta)} \leq c := \sup_{h \in \mathcal{H}} \frac{D\mathcal{J}(\sigma^0, h, \theta^*)}{\mathcal{J}(\sigma^0, \theta^*)}.$$

This characterization of an optimal strategy θ^* of the problem (3.26) leads to the

Definition 3.2. The trading strategy $\theta^* \in \Theta$ is called optimal mean-variance robust trading strategy against the class of alternatives \mathcal{H} if it is a solution of the optimization problem

$$\begin{aligned} & \text{minimize } \mathcal{J}(\sigma^0, \theta) \text{ over all strategies } \theta \in \Theta, \text{ subject to constraint} \\ & \sup_{h \in \mathcal{H}} \frac{D\mathcal{J}(\sigma^0, h, \theta)}{\mathcal{J}(\sigma^0, \theta)} \leq c, \end{aligned} \quad (3.27)$$

where c is some generic constant.

Remark 3.2. In contrast to “mean-variance robust” trading strategy which associates with optimization problem (3.25) and control theory, we find the “optimal mean-variance robust” strategy in the sense of Definition 3.2. Such approach and term are common in robust statistics theory (see, e.g., Hampel et al. [12], Rieder [34]).

Does the suggested approach provide “good” approximation? Consider the case.

Diffusion model with zero drift. Let a standard Wiener process $w = (w_t)_{0 \leq t \leq T}$ be given on the complete probability space (Ω, \mathcal{F}, P) . Denote by $F^w = (\mathcal{F}_t^w, 0 \leq t \leq T)$ the P -augmentation of the natural filtration $\mathcal{F}_t^w = \sigma(w_s, 0 \leq s \leq t), 0 \leq t \leq T$, generated by w .

Let the stock price process be modeled by the equation

$$dX_t^\sigma = X_t^\sigma \cdot \sigma_t dw_t, \quad X_0^\sigma > 0, \quad 0 \leq t \leq T,$$

where $\sigma \in A_\delta$ with

$$\int_0^T (\sigma_t^0)^2 dt < \infty$$

and $h \in \text{Ball}_{L_\infty(dt \times dP)}(0, r)$, $0 < r < \infty$. All considered processes are real-valued.

Denote by R^σ the yield process, i.e.,

$$dR_t^\sigma = \sigma_t dw_t, \quad R_0^\sigma = 0, \quad 0 \leq t \leq T.$$

The wealth at maturity T , with the initial endowment x , is equal to

$$V_T^{x,\theta}(\sigma) = x + \int_0^T \theta_t dR_t^\sigma.$$

Let, further, the contingent claim H be \mathcal{F}_T^w -measurable P -square-integrable r.v.

Consider the optimization problem (3.25). It is easy to see that if $\sigma \in A_\delta$; then

$$\sigma_t^0 - \delta r \leq \sigma_t \leq \sigma_t^0 + \delta r, \quad 0 \leq t \leq T, \quad P\text{-a.s.},$$

By the martingale representation theorem

$$H = EH + \int_0^T \varphi_t^H dw_t, \quad P\text{-a.s.},$$

where φ^H is the F^w -predictable process with

$$E \int_0^T (\varphi_t^H)^2 dt < \infty. \tag{3.28}$$

Hence

$$E(H - V_T^{x,\theta}(\sigma))^2 = (EH - x)^2 + E \int_0^T (\varphi_t^H - \sigma_t \theta_t)^2 dt.$$

From this it directly follows that the process

$$\begin{aligned} \sigma_t^*(\theta) &= (\sigma_t^0 - \delta r) I_{\{\frac{\varphi_t^H}{\theta_t} \geq \sigma_t^0\}} I_{\{\theta_t \neq 0\}} \\ &\quad + (\sigma_t^0 + \delta r) I_{\{\frac{\varphi_t^H}{\theta_t} < \sigma_t^0\}} I_{\{\theta_t \neq 0\}}, \quad 0 \leq t \leq T, \end{aligned} \tag{3.29}$$

is a solution of the optimization problem

$$\text{maximize } E(H - V_T^{x,\theta}(\sigma))^2 \text{ over all } \sigma \in A_\delta, \text{ with a given } \theta \in \Theta.$$

It remains to minimize (w.r.t. θ) the expression

$$E \int_0^T (\varphi_t^H - \sigma_t^*(\theta)\theta_t)^2 dt.$$

From (3.29) it easily follows that the equation (w.r.t. θ)

$$\varphi_t^H - \sigma_t^*(\theta)\theta_t = 0,$$

has no solution, but

$$\theta_t^* = \frac{\varphi_t^H}{\sigma_t^0} I_{\{\sigma_t^0 \neq 0\}}, \quad 0 \leq t \leq T, \quad (3.30)$$

solves problem. We assume that $0/0 := 0$.

Consider now the optimization problem (3.27).

For each fixed h

$$\begin{aligned} J(\sigma, \theta) &= E \left(H - x - \int_0^T \theta_t dR_t^\sigma \right)^2 \\ &= E \left(H - x - \int_0^T \theta_t \sigma_t^0 dw_t - \delta \int_0^T \theta_t h_t dw_t \right)^2 \\ &= J(\sigma^0, \theta) - 2\delta E \left[\left(EH - x + \int_0^T (\varphi_t^H - \theta_t \sigma_t^0) dw_t \right) \int_0^T \theta_t h_t dw_t \right] \\ &\quad + \delta^2 E \int_0^T \theta_t^2 h_t^2 dt, \end{aligned}$$

and hence

$$DJ(\sigma^0, h; \theta) = 2E \int_0^T (\theta_t \sigma_t^0 - \varphi_t^H) \theta_t h_t dt, \quad (3.31)$$

as follows from (3.28), the definition of the class \mathcal{H} and the estimation

$$\begin{aligned} \left(E \int_0^T (\theta_t \sigma_t^0 - \varphi_t^H) \theta_t h_t dt \right)^2 &\leq E \int_0^T (\theta_t \sigma_t^0 - \varphi_t^H)^2 dt E \int_0^T \theta_t^2 h_t^2 dt \\ &\leq \text{const} \cdot r^2 \left(E \int_0^T \theta_t^2 (\sigma_t^0)^2 dt + E \int_0^T (\varphi_t^H)^2 dt \right) E \int_0^T \theta_t^2 dt < \infty. \end{aligned} \quad (3.32)$$

Since, further, $DJ(\sigma^0, h; \theta) = 0$ for $h \equiv 0$, using (3.32) we get

$$0 \leq \sup_{h \in \mathcal{H}} DJ(\sigma^0, h; \theta) < \infty.$$

Hence we can take $0 \leq c < \infty$ in problem (6). Now if we substitute θ^* from (3.30) into (3.31), we get $DJ(\sigma^0, h; \theta^*) = 0$ for each h , and thus

$$\frac{\sup_{h \in \mathcal{H}} DJ(\sigma^0, h; \theta^*)}{J(\sigma^0, \theta^*)} = 0.$$

If we recall that $\theta^* = \arg \min_{\theta \in \Theta_{A_\delta}} J(\sigma^0, \theta)$, we get that θ^* defined by (3.30) is a solution of this optimization problem as well.

Thus we prove that

(a) *the mean-variance robust trading strategy $\theta^* = (\theta_t^*)_{0 \leq t \leq T}$ for the optimization problem (3.25) is given by the formula*

$$\theta_t^* = \frac{\varphi_t^H}{\sigma_t^0} I_{\{\sigma_t^0 \neq 0\}};$$

(b) *at the same time this strategy is an optimal mean-variance robust trading strategy for the optimization problem (3.27).*

Hence in this case the suggested approach leads to the perfect solution of initial problem (3.25).

To solve the problem (3.27) in general case we need to calculate $DJ(\sigma^0, h, \theta)$. Suppose that $k = (k_t)_{0 \leq t \leq T} = (k_{i,t}, 1 \leq i \leq d)_{0 \leq t \leq T}$ from (3.10) is such that $|k_{i,t}| \leq \text{const } \forall i, t$.

Following RSch [33] and GLP [11] introduce the probability measure $\tilde{Q} \sim P$ on \mathcal{F}_T by the relation

$$d\tilde{Q} = \frac{\tilde{z}_T}{\tilde{z}_0} d\tilde{P} \quad \left(\text{and hence } d\tilde{Q} = \frac{\tilde{z}_T^2}{\tilde{z}_0} dP \right). \quad (3.33)$$

Using Proposition 5.1 of GLP [11] we can write

$$\begin{aligned} \mathcal{J}(\sigma, \theta) &= E \frac{\tilde{z}_T^2}{\tilde{z}_0^2} \frac{\tilde{z}_0^2}{\tilde{z}_T^2} \left(H - x - \int_0^T \theta'_t dR_t^\sigma \right)^2 \\ &= \tilde{z}_0^{-1} E^{\tilde{Q}} \frac{\tilde{z}_0^2}{\tilde{z}_T^2} \left(H - x - \int_0^T \theta'_t \sigma_t dM_t^0 \right)^2 \\ &= \tilde{z}_0^{-1} E^{\tilde{Q}} \left(\frac{H \tilde{z}_0}{\tilde{z}_T} - x - \int_0^T \psi_t^0(\sigma) d \frac{\tilde{z}_0^2}{\tilde{z}_t^2} - \int_0^T (\psi_t^1(\sigma))' d \frac{M_t^0}{\tilde{z}_t} \tilde{z}_0 \right)^2 \\ &:= \overline{\mathcal{J}}(\sigma, \psi^0, \psi^1) \quad (\text{or } \overline{\mathcal{J}}(\sigma, \psi) \text{ with } \psi = (\psi^0, \psi^1)'), \end{aligned} \quad (3.34)$$

where

$$\begin{aligned} \psi_t^1 &= \psi_t^1(\sigma) = \sigma'_t \theta_t, \\ \psi_t^0 &= \psi_t^0(\sigma) = \int_0^t \theta'_s \sigma_s dM_s^0 - \theta'_t \sigma_t M_t^0, \quad 0 \leq t \leq T. \end{aligned} \quad (3.35)$$

Thus

$$\psi_t^1(\sigma) = \psi_t^1(\sigma^0) + \delta\psi_t^1(h), \quad \psi_t^0(\sigma) = \psi_t^0(\sigma^0) + \delta\psi_t^0(h).$$

Let (following RSch [33])

$$\frac{H}{\tilde{z}_T} \tilde{z}_0 = E \left(\frac{H}{\tilde{z}_T} \tilde{z}_0 \right) + \int_0^T (\psi_t^H)' dU_t + L_T, \quad (3.36)$$

be the Galtchouk–Kunita–Watanabe decomposition of r.v. $\frac{H}{\tilde{z}_T} \tilde{z}_0$ w.r.t. $R^{(d+1)}$ -valued \tilde{Q} -local martingale $U = (\frac{\tilde{z}_0}{\tilde{z}}, \frac{M^0}{\tilde{z}} \tilde{z}_0)'$, where $\psi^H = (\psi^{0,H}, \psi^{1,H})' \in L^2(U, \tilde{Q})$, the space of F -predictable processes ψ such that $\int \psi' dU \in \mathcal{M}^2(\tilde{Q})$ of martingale, and $L \in \mathcal{M}_{0,\text{loc}}^2(\tilde{Q})$, L is \tilde{Q} -strongly orthogonal to U .

Remember that

$$\psi = (\psi^0, \psi^1)'. \quad (3.37)$$

Then, using (3.34), (3.35) and (3.36) we can write for each h

$$\begin{aligned} & \mathcal{J}(\sigma^0 + \delta h, \psi) = \mathcal{J}(\sigma^0, \psi) + \delta \cdot 2\tilde{z}_0^{-1} \\ & \times E^{\tilde{Q}} \left\{ \left[\left(x - E^{\tilde{Q}} \frac{H}{\tilde{z}_T} \tilde{z}_0 \right) - L_T + \int_0^T (\psi_t(\sigma^0) - \psi_t^H)' dU_t \right] \int_0^T (\bar{\psi}_t(h))' dU_t \right\} \\ & + \delta^2 \tilde{z}_0^{-1} E^{\tilde{Q}} \left[\int_0^T (\bar{\psi}_t(h))' dU_t \right]^2 \\ & = \mathcal{J}(\sigma^0, \psi) + \delta \cdot 2\tilde{z}_0^{-1} E^{\tilde{Q}} \left[\int_0^T (\psi_t(\sigma^0) - \psi_t^H)' dU_t \int_0^T (\bar{\psi}_t(h))' dU_t \right] \\ & + \delta^2 \tilde{z}_0^{-1} E^{\tilde{Q}} \left[\int_0^T (\bar{\psi}_t(h))' dU_t \right]^2. \end{aligned} \quad (3.38)$$

Using Proposition 8 of RSch [33] we have for each h

$$\frac{\tilde{z}_0}{\tilde{z}_T} G_r(h, \Theta) = \left\{ \int_0^T (\psi(h))' dU_t : \psi(h) \in L^2(U, \tilde{Q}) \right\},$$

and hence by (3.23)

$$\begin{aligned} & E^{\tilde{Q}} \left(\int_0^T (\psi_t(h))' dU_t \right)^2 \\ & = E^{\tilde{Q}} \frac{\tilde{z}_0^2}{\tilde{z}_T^2} G_T^2(h, \theta) = \tilde{z}_0 E G_T^2(h, \theta) = \tilde{z}_0 E \left(\int_0^T \theta_t' dR_t^h \right)^2 \end{aligned}$$

$$\begin{aligned}
&= \tilde{z}_0 E \left(\int_0^T \theta'_t h_t dM_t^0 \right)^2 = \tilde{z}_0 E \left(\int_0^T \theta'_t h_t d\langle M \rangle_t k_t + \int_0^T \theta'_t h_t dM_t \right)^2 \\
&\leq \text{const} \left[E \left(\int_0^T |\theta'_t h_t d\langle M \rangle_t k_t| \right)^2 + E \left(\int_0^T \theta'_t h_t dM_t \right)^2 \right] \\
&\leq \text{const } r^2 E \int_0^T |\theta_t|^2 dC_t < \infty.
\end{aligned} \tag{3.39}$$

Further,

$$\begin{aligned}
&\left(E^{\tilde{Q}} \left[\int_0^T (\psi_t(\sigma^0) - \psi_t^H)' dU_t \int_0^T (\psi_t(h))' dU_t \right] \right)^2 \\
&\leq E^{\tilde{Q}} \left(\int_0^T (\psi_t(\sigma^0) - \psi_t^H)' dU_t \right)^2 E^{\tilde{Q}} \left(\int_0^T (\psi_t(h))' dU_t \right)^2 < \infty.
\end{aligned} \tag{3.40}$$

From these estimates we conclude that:

$$1) \quad D\overline{\mathcal{J}}(\sigma^0, h, \psi) = 2\tilde{z}_0^{-1} E^{\tilde{Q}} \int_0^T (\psi_t(\sigma^0) - \psi_t^H)' d\langle U \rangle_t \psi_t(h) < \infty, \tag{3.41}$$

thanks to (3.39).

2) $D\overline{\mathcal{J}}(\sigma^0, h, \psi)|_{h \equiv 0} = 0$, since $\psi(0) = 0$ by (3.37) and (3.35).

Thus

$$\sup_{h \in \mathcal{H}} D\overline{\mathcal{J}}(\sigma^0, h, \psi) \geq 0. \tag{3.42}$$

3) From (3.40) and (3.39) we get

$$\begin{aligned}
&(D\overline{\mathcal{J}}(\sigma^0, h, \psi))^2 \leq \text{const } \tilde{z}_0^{-2} r^2 \\
&\times E^{\tilde{Q}} \int_0^T (\psi_t(\sigma^0) - \psi_t^H)' d\langle U \rangle_t (\psi_t(\sigma^0) - \psi_t^H) E \int_0^T |\theta_t|^2 dC_t < \infty.
\end{aligned}$$

Thus $|D\overline{\mathcal{J}}(\sigma^0, h, \psi)|$ is estimated by the expression which does not depend on h , and is equal to zero if we substitute $\psi_t(\sigma^0) \equiv \psi_t^H$, $0 \leq t \leq T$.

Hence, by (3.42)

$$0 \leq \sup_{h \in \mathcal{H}} D\overline{\mathcal{J}}(\sigma^0, h, \psi)|_{\psi \equiv \psi^H} \leq \sup_{h \in \mathcal{H}} |D\overline{\mathcal{J}}(\sigma^0, h, \psi)|_{\psi \equiv \psi^H} = 0 \tag{3.43}$$

Further, from (3.42) follows that we can take $c \in [0, \infty)$ in (3.27).

Now substituting $\psi \equiv \psi^H$ into $\overline{\mathcal{J}}(\sigma^0, \psi)$ and $D\overline{\mathcal{J}}(\sigma^0, h, \psi)$ we get

$$\overline{\mathcal{J}}(\sigma^0, \psi^H) = \min_{\psi} \overline{\mathcal{J}}(\sigma^0, \psi) = \tilde{z}_0^{-1} (E^{\tilde{P}} H - x)^2 + \tilde{z}_0^{-1} E^{\tilde{Q}} L_T^2$$

(see Lemma 5.1 of GLP [11]) and

$$\sup_{h \in \mathcal{H}} \frac{D\overline{\mathcal{J}}(\sigma^0, h, \psi^H)}{\overline{\mathcal{J}}(\sigma^0, \psi^H)} = 0.$$

Hence the constraint of problem (3.27) is satisfied.

Remark 3.3. If $x = E^{\tilde{P}}H$ and $L_T \equiv 0$, then we get

$$\frac{D\overline{\mathcal{J}}(\sigma^0, h, \psi^H)}{\overline{\mathcal{J}}(\sigma^0, \psi^H)} = \frac{0}{0}$$

which is assumed to be zero, since if we consider the shifted risk functional $\tilde{\mathcal{J}} = \overline{\mathcal{J}} + 1$, the optimization problem and the optimal trading strategy will not change, but $D\tilde{\mathcal{J}}(\sigma^0, h, \psi^H) = D\overline{\mathcal{J}}(\sigma^0, h, \psi^H) = 0$ and $\tilde{\mathcal{J}}(\sigma^0, \psi^H) = 1$.

Finally, using Proposition 8 of RSch [33] we arrive at the following

Theorem 3.1. *In Model (3.10) under conditions (c.1) and (c.2) the optimal mean-variance robust trading strategy (in the sense of Definition 3.1) is given by the formula*

$$\theta_t^* = ((\sigma_t^0)')^{-1}[\psi_t^{1,H} + \zeta_t(V_t^* - (\psi_t^H)'U_t)], \quad 0 \leq t \leq T, \quad (3.44)$$

where

$$\begin{aligned} \psi_t^H &= (\psi_t^{0,H}, \psi_t^{1,H})', \quad U_t = \left(\frac{\tilde{z}_0}{\tilde{z}_t}, \frac{M_t^0}{\tilde{z}_t} \tilde{z}_0 \right)', \\ V_t^* &= \frac{\tilde{z}_0}{\tilde{z}_t} \left(x + \int_0^t (\psi_s^H)' dU_s \right), \end{aligned}$$

ψ_t^H and ζ_t are given by the relations (3.36) and (3.11), respectively, \tilde{z}_t is defined in (3.11).

REFERENCES

- [1] M. Avellaneda, A. Levy and A. Paras, Pricing and hedging derivative securities in markets with uncertain volatilities. *Appl. Math. Finance* **2** (1995), 73–88.
- [2] M. Avellaneda and A. Paras, Managing the volatility risk of portfolios of derivative securities: The Lagrangian uncertain volatility model. *Appl. Math. Finance* **3** (1996), 21–52.
- [3] F. Biagini, P. Guasoni and M. Pratelli, Mean-variance hedging for stochastic volatility models. *INFORMS Applied Probability Conference (Ulm, 1999)*. *Math. Finance* **10** (2000), No. 2, 109–123.
- [4] I. Bajeux-Besnainou and J. Rochet, Dynamic spanning: Are options an appropriate instrument? *Math. Finance* **6** (1996), No. 1, 1–16.
- [5] M. Chesney, R. J. Elliott, D. Madan and H. Yang, Diffusion coefficient estimation and asset pricing when risk premia and sensitivities are time varying. *Math. Finance* **3** (1993), No. 2, 85–99.
- [6] R. Chitashvili, N. Lazrieva and T. Toronjadze, Asymptotic theory of M -estimators in general statistical models. Part I. On asymptotic behaviour of estimators under model perturbation. *Centrum voor Wiskunde en Informatica Amsterdam*, Report BS-R9019, 1990.

- [7] R. Chitashvili, N. Lazrieva and T. Toronjadze, Asymptotic theory of M -estimators in general statistical models. Part II. On asymptotic behaviour of estimators in the presence of nuisance parameter. *Centrum voor Wiskunde en Informatica Amsterdam*, Report BS-R9020, 1990.
- [8] F. Delbaen, P. Monat, W. Schachermayer, M. Schweizer and C. Stricker, Weighted norm inequalities and hedging in incomplete markets. *Finance Stoch.* **1** (1997), No. 3, 181–227.
- [9] D. Duffie and H. R. Richardson, Mean-variance hedging in continuous time. *Ann. Appl. Probab.* **1** (1991), No. 1, 1–15.
- [10] H. Föllmer and D. Sondermann, Hedging of nonredundant contingent claims. *Contributions to mathematical economics*, 205–223, *North-Holland, Amsterdam*, 1986.
- [11] C. Gouieroux, J. P. Laurent and H. Pham, Mean-variance hedging and numéraire. *Math. Finance* **8** (1998), No. 3, 179–200.
- [12] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw and W. A. Stahel, Robust statistics. The approach based on influence functions. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. John Wiley & Sons, Inc., New York, 1986.
- [13] J. Hull and A. White, The pricing of options on assets with stochastic volatilities. *J. Finance* **3** (1987), 281–300.
- [14] I. A. Ibragimov, R. Z. Has'minskiĭ, Statistical estimation. Asymptotic theory. (Translated from the Russian) *Applications of Mathematics*, 16. Springer-Verlag, New York-Berlin, 1981.
- [15] J. Jacod and A. N. Shiryaev, Limit theorems for stochastic processes. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, 288. Springer-Verlag, Berlin, 1987.
- [16] N. V. Krylov, Controlled diffusion processes. (Translated from the Russian) *Applications of Mathematics*, 14. Springer-Verlag, New York-Berlin, 1980.
- [17] Yu. Kutoyants, Identification of dynamical systems with small noise. *Mathematics and its Applications*, 300. Kluwer Academic Publishers Group, Dordrecht, 1994.
- [18] J. P. Laurent and H. Pham, Dynamic programming and mean-variance hedging. *Finance Stoch.* **3** (1999), No. 1, 83–110.
- [19] N. L. Lazrieva and T. A. Torondzhadze, Asymptotic properties of an M -estimate in a general statistical experiment scheme. (Russian) *Statistics and control of random processes* (Russian) (*Preila*, 1987), 105–112, *Nauka, Moscow*, 1989.
- [20] N. Lazrieva and T. Toronjadze, Robust estimators in statistical models with filtration. Shrinking neighbourhoods. *Seminarberichte, Fachbereich Mathematik, Fernuniversität, Hagen, Germany* **48** (1994), 50–68.
- [21] N. Lazrieva and T. Toronjadze, Robust estimators in statistical models associated with semimartingales. *Proc. A. Razmadze Math. Inst.* **118** (1998), 73–100.
- [22] D. Li and W.-L. Ng, Optimal dynamic portfolio selection: multiperiod mean-variance formulation. *Math. Finance* **10** (2000), No. 3, 387–406.
- [23] R. S. Liptser and A. N. Shiryaev, Statistics of random processes. I. General theory. (Translated from the Russian) *Applications of Mathematics*, Vol. 5. Springer-Verlag, New York-Heidelberg, 1977.
- [24] M. Mania and R. Tevzadze, A semimartingale Bellman equation and the variance-optimal martingale measure. *Georgian Math. J.* **7** (2000), No. 4, 765–792.
- [25] M. Mania, M. Santacrose and R. Tevzadze, A semimartingale backward equation related to the p -optimal martingale measure and the lower price of a contingent claim. *Stochastic processes and related topics (Siegmundsborg, 2000)*, 189–212, *Stochastics Monogr.*, 12, Taylor & Francis, London, 2002.
- [26] H. M. Markowitz, Mean-variance analysis in portfolio choice and capital markets. *Basil Blackwell, Oxford*, 1987.

- [27] R. C. Merton, An analytic derivation of the efficient portfolio frontier. *J. Financ. Qualitat. Anal.* **7** (1972), 1851–1872.
- [28] P. Monat and C. Stricker, Föllmer-Schweizer decomposition and mean-variance hedging for general claims. *Ann. Probab.* **23** (1995), No. 2, 605–628.
- [29] M. Musiela and M. Rutkowski, Martingale methods in financial modelling. *Applications of Mathematics (New York)*, 36. Springer-Verlag, Berlin, 1997.
- [30] S. Pastorello, Diffusion coefficient estimation and asset pricing when risk premia and sensitivities are time varying: A comment. *Math. Finance* **6** (1996), No. 1, 111–117.
- [31] H. Pham, T. Rheinländer and M. Schweizer, Mean-variance hedging for continuous processes: new proofs and examples. *Finance Stoch.* **2** (1998), No. 2, 173–198.
- [32] E. Renault and N. Touzi, Option hedging and implied volatilities in a stochastic volatility model. *Math. Finance* **6** (1996), No. 3, 279–302.
- [33] T. Rheinländer and M. Schweizer, On L^2 -projections on a space of stochastic integrals. *Ann. Probab.* **25** (1997), No. 4, 1810–1831.
- [34] H. Rieder, Robust asymptotic statistics. *Springer Series in Statistics*. Springer-Verlag, New York, 1994.
- [35] W. J. Runggaldier and A. Zaccaria, A stochastic control approach to risk management under restricted information. *INFORMS Applied Probability Conference (Ulm, 1999)*. *Math. Finance* **10** (2000), No. 2, 277–288.
- [36] M. Schweizer, Mean-variance hedging for general claims. *Ann. Appl. Probab.* **2** (1992), No. 1, 171–179.
- [37] M. Schweizer, Approximating random variables by stochastic integrals. *Ann. Probab.* **22** (1994), No. 3, 1536–1575.
- [38] M. Schweizer, Approximation pricing and the variance-optimal martingale measure. *Ann. Probab.* **24** (1996), No. 1, 206–236.
- [39] L. Scott, Option pricing when the variance changes randomly: theory, estimation and application. *J. Financ. Qualitat. Anal.* **22** (1987), 419–438.
- [40] K. R. Sircar and G. G. Papanicolaou, Stochastic volatility, smile and asymptotics. *Appl. Math. Finance* **6** (1999), No. 2, 107–145.
- [41] T. Toronjadze, Optimal mean-variance robust hedging under asset price model misspecification. *Georgian Math. J.* **8** (2001), No. 1, 189–199.